(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 0 380 563 B1

(12) EUROPEAN PATENT SPECIFICATION

(45) Date of publication and mention of the grant of the patent:
09.12.1998 Bulletin 1998/50

(21) Application number: **88908903.3**

(22) Date of filing: **22.09.1988**

(51) Int Cl.[6]: **H04B 15/00**

(86) International application number:
PCT/US88/03269

(87) International publication number:
WO 89/03141 (06.04.1989 Gazette 1989/08)

(54) IMPROVED NOISE SUPPRESSION SYSTEM

LÄRMUNTERDRÜCKUNGSSYSTEM

SYSTEME AMELIORE DE SUPPRESSION DU BRUIT

(84) Designated Contracting States:
CH DE FR GB IT LI NL

(30) Priority: **01.10.1987 US 103857**

(43) Date of publication of application:
08.08.1990 Bulletin 1990/32

(73) Proprietor: **MOTOROLA, INC.**
Schaumburg, IL 60196 (US)

(72) Inventors:
- **VILMUR, Richard, Joseph**
  Palatine, IL 60067 (US)
- **BARLO, Joseph, John**
  Hoffman Estates, IL 60195 (US)
- **GERSON, Ira, Alan**
  Hoffman Estates, IL 60195 (US)
- **LINDSLEY, Brett, Louis**
  Palatine, IL 60067 (US)

(74) Representative: **Hudson, Peter David et al**
Motorola,
European Intellectual Property,
Midpoint,
Alencon Link
Basingstoke, Hampshire RG21 7PL (GB)

(56) References cited:

US-A- 4 110 784
US-A- 4 287 475
US-A- 4 325 068
US-A- 4 628 529
US-A- 4 630 304
US-A- 4 630 305
US-A- 4 635 217
US-A- 4 648 127

- **ICASSP 79, 1979 IEEE INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH, AND SIGNAL PROCESSING, Washington, D.C., 2nd - 4th April 1979, pages 200-203, IEEE, New York, US; S.F. BOLL: "A spectral subtraction algorithm for suppression of acoustic noise in speech"**
- **SIGNAL PROCESSING, vol. 8, no. 4, July 1985, pages 387-400, Elsevier Science Publishers B.V. (North Holland), Amsterdam, NL; P. VARY et al.: "Noise suppression by spectral magnitude estimation - mechanism and theoretical limits"**
- **1978 IEEE INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH & SIGNAL PROCESSING, Tulsa, Oklahoma, 10th - 12th April 1978, pages 602-605, IEEE, New York, US; R. A. CURTIS et al.: "An investigation of several frequency-domain processing methods for enhancing the intelligibility of speech in wideband random noise"**
- **ICASSP 87, 1987 INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH, AND SIGNAL PROCESSING, Dallas, Texas, 6th - 9th April 1987, vol. 2, pages 1143-1146, IEEE, New York, US; D. VAN COMPERNOLLE: "Increased noise immunity in large vocabulary speech recognition with the aid of spectral subtraction"**

Note: Within nine months from the publication of the mention of the grant of the European patent, any person may give notice to the European Patent Office of opposition to the European patent granted. Notice of opposition shall be filed in a written reasoned statement. It shall not be deemed to have been filed until the opposition fee has been paid. (Art. 99(1) European Patent Convention).

EP 0 380 563 B1

Printed by Jouve, 75001 PARIS (FR)

Description

Cross-Reference to Related Applications

This application includes reference to U.S. Patent No. 4,628,529, assigned to the same assignee as the present application. Furthermore, this application contains subject matter related to U.S. Patent No. 4,630,304 and U.S. Patent No. 4,630,305, also assigned to the same assignee as the present application.

Background of the Invention

1. Field of the Invention

The present invention relates generally to acoustic noise suppression systems. The present invention is more specifically directed to improving the speech quality of a noise suppression system employing the spectral subtraction noise suppression technique.

2. Description of the Prior Art

Acoustic noise suppression in a speech communication system generally serves the purpose of improving the overall quality of the desired audio signal by filtering environment background noise from the desired speech signal. This speech enhancement process is particularly necessary in environments having abnormally high levels of ambient background noise, such as an aircraft, a moving vehicle, or a noisy factory.

The noise suppression technique described in the aforementioned patents is the spectral subtraction - or spectral gain modification --technique. Using this approach, the audio input signal is divided into individual spectral bands by a bank of bandpass filters, and particular spectral bands are attenuated according to their noise energy content. A spectral subtraction noise suppression prefilter utilizes an estimate of the background noise power spectral density to generate a signal-to-noise ratio (SNR) of the speech in each channel, which, in turn, is used to compute a gain factor for each individual channel. The gain factor is used as a pointer for a look-up table to determine the attenuation for that particular spectral band. The channels are then attenuated and recombined to produce the noise-suppressed output waveform.

In specialized applications involving relatively high background noise environments, most noise suppression techniques exhibit significant performance limitations. One example of such an application is the vehicle speakerphone option to a cellular mobile radio telephone system, which provides hands-free operation for the automobile driver. The mobile hands-free microphone is typically located at a greater distance from the user, such as being mounted overhead on the visor. The more distant microphone delivers a much poorer signal-to-noise ratio to the land-end party drive to road and wind noise conditions. Although the received speech at the land-end is usually intelligible, continuous exposure to such background noise levels often increases listener fatigue.

Although most prior art techniques perform sufficiently well under nominal background noise conditions, the performance of known techniques becomes severely limited in such specialized applications of unusually high background noise. Typical spectral subtraction noise suppression systems may reduce the background noise level over the voice frequency spectrum by as much as 10 dB without seriously affecting the speech quality. However, when the prior art techniques are used in relatively high background noise environments requiring noise suppression levels approaching 20 dB, there is a substantial degradation in the quality characteristics of the voice. Furthermore, in rapidly-changing high noise environments, a severe low frequency noise flutter develops in the output speech signal which resembles a distant "jet engine roar" sound. This noise flutter is inherent in as spectral subtraction noise suppression system, since the individual channel gain parameters are continuously being updated in response to the changing background noise environment.

The background noise flutter problem was indirectly addressed but not eliminated through the use of gain smoothing. For example, R.J. McAulay and M.L. Malpass, in the article entitled "Speech Enhancement Using a soft-Decision Noise Suppression Filter", IEEE Trans. Amust.. Speech, Signal Processing, Vol. ASSP-28, No.2 (April 1980), pp. 137-145, propose the use of gain smoothing on a per-frame basis to avoid the introduction of discontinuities in the output waveform. Since the introduction of gain smoothing can cause the noise suppression prefilter to be slow to respond to a leading edge transition (which would result in speech distortion), a weighting a factor of 1 or 1/2 was chosen such that the prefilter responds immediately to an increase in gain while tending to smooth any decrease in gain. Unfortunately, excessive gain smoothing still produces noticeable detrimental effects in voice quality, the primary effect being the apparent introduction of a tail-end echo or "noise pump" to spoken words. There is also a significant reduction in voice amplitude with large amounts of gain smoothing.

The noise flutter performance was further improved by the technique of smoothing the noise suppression gain

factors for each individual channel on a per-sample basis instead of on a per-frame basis. Per-sample smoothing, as well as utilizing different smoothing coefficients for each channel, is described in U.S. Patent No. 4,630,305, entitles "Automatic Gain Selector for a Noise Suppression System." However, none of the known prior art techniques appreciate that the primary source of the channel gain discontinuities is the inherent fluctuation of background noise in each channel from one frame to the next. In known spectral subtraction systems, even a 2 dB SNR variation would create a few dB of gain variation, which is then heard as an annoying background noise flutter. Hence, the flutter problem has never been effectively solved.

Moreover, narrowband noise – that which has a high power spectral density in only a few channels – further complicates the background noise flutter problem. Since these few high energy noise channels would not be attenuated by the background noise suppression, the resultant audio output has a "running water" type of characteristic. Narrowband noise bursts also degrade the accuracy of the background noise update decision required to perform noise suppression in changing background noise environments.

Since the gain factors are chosen by SNR estimates, which are determined by the speech energy in each channel (signal) and the current background noise energy estimate in each channel (noise), the performance of the entire noise suppression system is based upon the accuracy of the background noise estimate. The statistics of the background noise are estimated during the time when only background noise is present, such as during the pause in human speech. Therefore, an accurate speech/noise clarification must be made to determine when such pauses in speech are occurring.

It is widely known that the energy histogram technique for distinguishing between background noise and speech perform sufficiently well in normal ambient noise environments. See, e.g., W.J. Hess, "A Pitch Synchronous Digital Feature Extraction system for Phonemic Recognition of Speech," IEEE Trans, Acoust., Speech, Signal Processing, Vol. ASSP-24, No. 1 (February 1976), pp. 14-25. Energy histograms of acoustic signals exhibit a bimodal distribution in which the two modes correspond to noise and speech. Thus, an appropriate threshold can be set between the two modes to provide the speech/noise classification. However, the distinction between background noise energy and unvoiced speech energy in relatively high background noise environments is unclear. Consequently, the task of accurately finding the two modes of the energy histogram, and setting the appropriate threshold between them , is extremely difficult.

To accommodate changing noise backgrounds, McAulay and Malpass implement and adaptive threshold by constantly monitoring the histogram energy on a frame-by-frame basis, and updating the threshold utilizing different decay factors. Alternatively, U.S. Patent No. 4,630,304 utilizes an energy valley detector to perform the speech/noise decision based upon the post-processed signal energy -- signal energy available at the output of the noise suppression system -- to determine the detected speech minimum. Thus, the accuracy of the background noise estimate is improved since it is based upon a much cleaner speech signal.

However, neither prior art technique is properly responsive to a sudden, strong increase in background noise level. These background noise estimate updating decision processes interpret a sudden, loud noise level rise as speech, such that no updates are performed. The energy histogram or valley detector has a slow adaptation characteristic which will eventually adapt to the higher noise level. However, this adaptation characteristic does lead to incorrect noise updates on the weaker energy portions of speech. This erroneous decision significantly degrades the performance of the noise suppression system.

A need, therefore exists for an improved acoustic noise suppression system which addresses the problems of background noise fluctuations, narrowband noise bursts, and sudden background noise.

Summary of the Invention

According to a first aspect of the present invention there is provided a noise suppression system for attenuating the background noise from a noisy input signal to produce a noise-suppressed output signal, said noise suppression system comprising: means for separating the input signal into a plurality of pre-processed signals representative of selected frequency channels; means for generating estimates of the signal-plus-noise energy and the noise energy in each individual channel; and means for producing a gain value for each individual channel in response to said channel energy estimates; the system characterized in that: said gain values have a minimum gain value for each channel; said means for producing a gain value includes threshold means for allowing gain values above said minimum gain value to be produced only when said signal-plus-noise energy estimates exceed said noise energy estimates by a predetermined amount; and means for modifying the gain of each of said plurality of pre-processed signals in response to said gain values to provide a plurality of post-processed signals.

In another aspect of the present invention there is provided a method of attenuating the background noise from a noisy input signal to produce a noise-suppressed output signal in a noise suppression system comprising the steps of: separating the input signal into a plurality of pre-processed signals representative of a number N of selected frequency channels; generating an estimate of the energy in each individual channel; generating and storing an estimate

of the background noise power spectral density of said pre-processed signals; and generating an estimate of the signal-to-noise ratio (SNR) in each individual channel based upon said background noise estimates and said channel energy estimates; the method characterized by the steps of: producing a gain value for each individual channel in response to said channel SNR estimates, wherein said gain values have a range of minimal values; and wherein said gain values producing step includes the steps of: providing a predefined SNR threshold and comparing (said channel SNR estimates to said predefined SNR threshold such that channels having SNR estimates below said SNR threshold produce gain values within said minimal range; and modifying the gain of each of said plurality of pre-processed signals in response to said gain values to provide a plurality of post-processed signals.

The present invention advantageously provides an improved method and apparatus for suppressing background noise in high background noise environments without significantly degrading voice quality. Moreover, the present invention addresses the problem of background noise fluctuation without requiring large amounts of gain smoothing. Additionally, a preferred embodiment of the present invention provides a spectral subtraction noise suppression system that compensates for the detrimental effects of narrowband noise bursts. Furthermore, the present invention provides an improved background noise estimation mechanism that is not misled by low energy portions of speech, yet still provides correction for sudden, strong increases in background noise levels.

In the present invention, the improvements to the noise suppression system relate to the addition of an SNR threshold mechanism to eliminate minor gain fluctuations for low SNR conditions, a voice metric calculator for producing a more accurate background noise estimate update decision and a channel SNR modifier to suppress narrowband noise bursts.

More specifically, the first aspect of the present invention pertains to the addition of an SNR threshold mechanism for providing a predetermined SNR threshold which the channel SNR estimates must exceed before a gain value above a predefined minimum gain value can be produced. In a preferred embodiment, the SNR threshold is set at 2.25 dB SNR, such that minor background noise fluctuations do not create step discontinuities in the noise suppression gains.

In a preferred embodiment, the a voice metric calculator is utilized to perform the speech/noise classification for the background noise update decision using a two-step process. First, the raw SNR estimates are used to index a voice metric table to obtain voice metric values for each channel. A voice metric is a measurement of the overall voice-like characteristics of all of the channel energies. The individual voice channel metric values are summed to create a first multi-channel energy parameter, and then compared to a background noise update threshold. If the voice metric sum does not meet the threshold, the input frame in deemed to be noise, and a background noise update is performed. Secondly, the time since the occurrence of the previous background estimate update is constantly monitored. If too much time has passed since the last update, e.g. 1 second, then it is assumed that a substantial increase in noise has occurred, and a background noise update is performed regardless of whether it looks like a voice frame. This second test is based on the assumption that speech seldom contains continuous high energy levels in all channels for more than one second, which would be the case for a sudden, loud noise level increase. The voice metric algorithm incorporating the two-step decision process provides a very accurate background noise estimate update signal.

In another preferred embodiment, a channel SNR modifying mechanism provides a second multi-channel energy parameter in response to the number of upper-channel SNR estimates which exceed a predetermined energy threshold, e.g. 6 dB SNR. If only a few channels have an energy level above this threshold (such as would be the cased for a narrowband noise burst), the measured SNR for those particular channels would be reduced. Moreover, if the aforementioned voice metric sum is less than a metric threshold (which would indicate that the frame was noise), all channels are similarly reduced. This SNR modifying technique is based on the assumption that typical speech exhibits a majority of channels having signal-to-noise ratios of 6 dB or greater.

An exemplary embodiment of the present invention will now be described with reference to the accompanying drawings.

Brief Description of the Drawings

Figure 1 is a detailed block diagram illustrating a preferred embodiment of the improved noise suppression system according to the present invention;

Figure 2 is a graph representing voice metric values output as a function of SNR estimate index values input for the voice metric calculator block of Figure 1;

Figure 3 is a representative gain table graph illustrating the overall channel attenuation for particular groups of channels as a function of the SNR estimate; and

Figures 4a through 4f are flowcharts illustrating the specific sequence of operations performed in accordance with the practice of the preferred embodiment of the present invention;

Figures 5a and 5b are flowcharts illustrating the general sequence of operations performed in the prior art;

Figures 6a to 6d are detailed flowcharts illustrating specific sequences of operations shown in Figures 5a and 5b.

Detailed Description of the Preferred Embodiment

Figure 1 is a detailed block diagram of the preferred embodiment of the present invention. All the elements of Figure 1 having reference numerals less than 600 correspond to those of U.S. Patent No.4,628,529 - Borth et al., and are shown in Figures 5a and 5b and Figures 6a to 6d of the present application. The additional circuit components having reference numerals greater than 600 represent the improvements to the system, and will be subsequently described herein.

First, concerning the prior art Borth patent U.S. 4,628,529 (and now expressly referring to Figures 5a and 5b and Figures 6a to 6d), FIG. 5a/b is a flowchart illustrating the overall operation of the prior art noise suppression mechanism. This generalized flow diagram is subdivided into three functional blocks: noise suppression loop 604 - further described in detail in FIG. 6a; automatic gain selector 615 - described in more detail in FIG. 6b; and automatic background noise estimator 621 - illustrated in FIGS. 7c and 7d.

The operation of the improved noise suppression system begins with FIG. 5a at initialization block 601. When the system is first powered-up, no old background noise estimate exists in energy estimate storage register 585, and no noise energy history exists in energy valley detector 570. Consequently, during initialization 601, storage register 585 is preset with an initialization value representing a background noise estimate value corresponding to a clean speech signal at the input. Similarly, energy valley detector 570 is preset with an initialization value representing a valley level corresponding to a noisy speech signal at the input.

Initialization block 601 also provides initial sample counts, channel counts, and frame counts. For the purposes of the following discussion, a sample period is defined as 125 microseconds corresponding to an 8 Khz sampling rate. The frame period is defined as being a 10 millisecond duration time interval to which the input signal samples are quantized. Thus, a frame corresponds to 80 samples at an 8 Khz sampling rate.

Initially, the sample count is set to zero. Block 602 increments the sample count by one, and a noisy speech sample is input from A/D converter 510 in block 603. The speech sample is then pre-emphasized by pre-emphasis network 520 in block 605.

Following pre-emphasis, block 606 initializes the channel count to one. Decision block 607 then tests the channel count number. If the channel count is less than the highest channel number N, the sample for that channel is bandpass filtered, and the signal energy for that channel is estimated in block 608. The result is saved for later use. Block 609 smoothes the raw channel gain for the present channel, and block 610 modifies the level of the bandpass-filtered sample utilizing the smoothed channel gain. The N Channels are then combined (also in block 610) to form a single processed output speech sample. Block 611 increments the channel count by one and the procedure in blocks 607 through 611 is repeated.

If the result of the decision in 607 is true, the combined sample is de-emphasized in block 612 and output as a modified speech sample in block 613. The sample count is then tested in block 614 to see if all samples in the current frame have been processed. If samples remain, the loop consisting of blocks 602 through 613 is re-entered for another sample. If all samples in the current frame have been processed, block 614 initiates the procedure if block 615 for updating the individual channel gains.

Continuing with FIG. 5b, block 616 initiates the channel counter to one. Block 617 tests if all channels have been processed. If this decision is negative, block 618 calculates the index to the gain table for the particular channel by forming an SNR estimate. This index is then utilized in block 619 to obtain a channel gain value from the look-up table. The gain value is then stored for use in noise suppression loop 604. Block 620 then increments the channel counter, and block 617 rechecks to see if all channel gains have been updated. If this decision is affirmative, the background noise estimate is then updated in block 621.

To update the background noise estimate, US-A-4,628,529 first simulates post-processed energy in block 622 by multiplying the updated raw channel gain value by the pre-processed energy estimate for that channel. Next, the simualted post-processed energy estimates are combined in block 623 to form an overall channel energy estimate for use by the valley detector. Block 624 compares the value of this overall post-processed energy estimate to the previous valley level. If the energy value exceeds the previous valley level, the previous valley level is updated in block 626 by increasing the level with a slow time constant. This occurs when voice, or a higher background noise level, is present. If the output of decision block 624 is negative (post-processed energy less than previous valley level), the previous valley level is updated in block 625 by decreasing the level with a fast time constant. This previous valley level decrease occurs when minimal background noise is present. Accordingly, the background noise history is continually updated by slowly increasing or rapidly decreasing the previous valley level towards the current post-processed energy estimate.

Subsequent to the updating of the previous valley level (block 625 or 626), decision block 627 tests if the current post-processed energy value exceeds a predetermined noise threshold. If the result of this comparison is negative, a decision that only noise is present is made, and the background noise spectral estimate is updated in block 628. This corresponds to the closing of channel switch 575. If the result of the test is affirmative, indicating that speech is present, the background noise estimate is not updated. In either case, the operation of background noise estimator 621 ends

when the sample count is rest in block 629 and the frame count is incremented in block 630. Operation then proceeds to block 602 to begin noise suppression on the next frame of speech.

The flowchart of FIG. 6a illustrates the specific details of the sequence of operation of noise suppression loop 604. For every sample of input speech, block 701 pre-emphasizes the sample by implement the filter described by the equation:

$$Y(nT)=X(nT)-K_1[X((n-1)T)]$$

where Y(nT) is the output of the filter at time nT, T is the sample period, X(nT) and X((n-1)T) are the input samples at times nT and (n-1)T respectively, and the pre-emphasis coefficient K_1 is 0.9375. As previously noted, this filter pre-emphasizes the speech sample at approximately +6 dB per octave.

Block 702 sets the channel count equal to one, and initializes the output sample total to zero. Block 703 tests to see is the channel count is equal to the total number of channels N. If this decision is negative, the noise suppression loop begins by filtering the speech sample through the bandpass filter corresponding to the present channel count. As noted earlier, the bandpass filters are digitally implemented using DSP techniques such that they function as 4-pole Butterworth bandpass filters.

The speech sample output from bandpass filter(cc) is then full-wave rectified in block 705, and low-pass filtered in block 706, to obtain the energy envelope value $E_{(cc)}$ for this particular sample. This channel energy estimate is then stored by block 707 for later use. As will be apparent to those skilled in the art, energy envelope value $E_{(cc)}$ is actually an estimate of the square root of the energy in the channel.

Block 708 obtains the raw gain value RG for channel cc and performs gain smoothing by means of a first order IIR filter, implementing the equation:

$$G(nT)=G((n-1)T)+K_2(cc)(RG(nT)-G(n-1)T)$$

where G(nT) is the smoothed channel gain at time nT, T is the sample period, G((n-1)T) is the smoothed channel gain at time (n-1)T, RG(nT) is the computed raw channel gain for the last frame period, and $K_2(cc)$ is the filter coefficient for channel cc. This smoothing of the raw gain values on a per-sample basis reduces the discontinuities in gain changes, thereby significantly improving noise flutter performance.

Block 709 multiplies the filtered sample obtained in block 704 by the smoothed gain value for channel cc obtained from block 708. This operation modifies the level of the bandpass filtered sample using the current channel gain, corresponding to the operation of channel gain modifier 250. Block 710 then adds the modified filter sample for channel cc to the output sample total, which, when performed N times, combines the N modified bandpass filter outputs to form a single processed speech sample output. The operation of block 710 corresponds to channel combined 260. Block 711 increments the channel count by one and the procedure in blocks 703 through 711 is then repeated.

If the result of the test in 703 is true, the output speech sample is de-emphasized at approximately -6 dB peroctave in block 712 according to the equation:

$$Y(nT)=X(nT)+K_3[Y((n-1)T)]$$

where X(nT) is the processed sample at time nT, T is the sample period, Y(nT) and Y((n-1)T) are the de-emphasized speech samples at times nT and (n-1)T respectively, and K_3 is the de-emphasis coefficient which has a value of 0.9375. The de-emphasized processed speech sample is then output to the D/A converter block 613. Thus, the noise suppression loop of FIG. 6a illustrates both the channel filter-bank noise suppression technique and the per-sample channel gain smoothing technique.

The flowchart of FIG. 6b more rigorously describes the- detailed detailed operation of automatic gain selector block 614 of FIG. 5. Following processing of all speech samples in a particular frame, the operation is turned over to block 615 which serves to update the individaul channel gains. First of all, the channel count (cc) is set to one in block 720. Next, decsion block 721 tests if all channels have been processed. If not, operation proceeds with block 722 which calculates the signal-to-noise ratio for the particulal channel. As previously mentioned, the SNR calculation is simply a division of the per-channel energy estimates (signal-plus-noise) by the perchannel background noise estimates (noise). Therefore, block 722 simply divides the current stored channel energy estimate from block 707 by the current background noise etimate from block 628 according to the equation:

Index (cc)=[current frame energy for channel cc]/background

noise energy estimate for channel cc].

In block 723, the particular gain table to be indexed is chosen. The quantized value of the current valley level is used to perform is selection. However, any method of gain table selection may be used. Furthermore, no gain table selection is required for noise suppression systems implementing a single gain table.

The SNR index calculated in block 722 is used in block 724 to look up the raw channel gain value from the appropriate gain table. Hence, the gain value is indexed as a function if two or three variables: (1) the channel number; (2) the current channel SNR estimate; and possibly (3) the overall average background noise level.

Block 725 stores the raw gain value chosen by block 724. The channel count is incremented in block 726, and then decision block 721 is re-entered. After all N channels gains have been updated, operation proceeds to block 621. Hence, automatic gain selector block 615 updates the channel gain values on a frame-by-frame basis to more accurately reflect the current SNR of each particular channel.

FIG. 6c and FIG. 6d expands upon block 621 to more specifically describe the function of a prior art automatic background noise estimator. Particularly, FIG. 6c describes the process of simulating the post-processed energy and combining these estimates, while FIG. 6d describes the operation of valley detector 570.

Referring now to FIG. 6c, the operation for simulating post-processed speech begins at block 730 by setting the channel count (cc) to one. Block 731 tests this channel count to see if all N channels have been processed. If not, the equation of block 732 describes the actual simulation process performed by a prior art energy estimate modifier.

Simulated post-processed speech energy is generated by multiplying the raw channel gain values (obtained directly from the channel gain tables) by the pre-processed energy estimate (obtained from channel energy estimator 220) for each channel via the equation:

$$SE(cc)=E(cc)\,RG(cc)$$

where SE(cc) is the simulated post-processed energy for channel cc, E(cc) is the current frame energy estimate for channel cc stored by block 707, and RG(cc) is the raw channel gain value for channel cc obtained from block 725. As noted earlier, E(cc) is actually the square root of the energy in the channel since it is a measure of the signal envelope. Hence, the RG(cc) term of the above equation is not squared. The multiplication performed in block 732 serves essentially the same function as channel gain modifier 250 - except that the channel gain modifier utilizes pre-processed speech *signal* whereas energy estimate modifier 560 utilizes pre-processed speech energy.

The channel counter is then incremented in block 733, and retested in block 731. When a simulated post-processed energy value is obtained for all N Channels, blocks 734 through 738 serve to combine the individual simulated channel energy estimates to form the single overall energy estimate according to the equation:

$$\sum_{i=1}^{N} \text{CHANNEL (i) POST-PROCESSED ENERGY}$$

where N is the number of filters in the filter bank. Block 734 initializes the channel count to one, and block 735 initializes the overall post-processed energy value to zero. After initialization, decision block 736 tests whether or not all channel energies have been combined. If not, block 737 adds the simulated post-processed energy value for the current channel to the overall post-processed energy value. The current channel number is then incremented in block 738, and the channel number is again tested at block 736. When all N channels have been combined to form the overall simulated post-processed energy estimate, operation proceeds to block 740 of FIG. 6d.

Referring now to FIG. 6d, blocks 740 through 745 illustrate how the post-processed signal energy is used to generate and update the previous valley level, corresponding to the operation of a prior art energy valley detector. After all the post-processed energies per channel have been combined, block 740 computes the logarithm of this combined post-processed channel energy. One reason that the log representation of the post processed speech energy is used in the embodiment is to facilitate implementation of an extremely large dynamic range (>90 dB) signal in an 8-bit microprocessor system.

Decision block 741 then tests to see if this log energy value exceeds the previous valley level. As previously mentioned, the previous valley level is either the stored valley level for the prior frame or an initialized valley level

provided by block 601 of FIG. 6. If the log value exceeds the previous valley level, the previous valley level is updated in block 743 with the current log [post-processed energy] value by increasing the level with the slow time constant of approximately one second to form a current valley level. This occurs when voice or a higher background noise level is present. Conversely, if the output of decision block 741 is negative (log[post-processed energy] less than previous valley level), the previous valley level is updated in block 742 with the current log [post-processed energy] value by decreasing the level with a fast time constant of approximately 40 milliseconds to form the current valley level. This occurs when a lower background noise level is present. Accordingly, the background noise history is continuously updated by slowly increasing or rapidly decreasing the previous valley level, depending upon the background noise level of the current simulated post-processed speech energy estimate.

After updating the previous valley level, decision block 744 tests if the current log [post-processed energy] value exceeds the current valley level plus a pre-determined offset. The addition of the current valley level plus this valley offset produces a noise threshold level. This offset provides approximately a 6 dB increase to the currant valley level. Hence, another reason for utilizing log arithmetic is to simplify the constant 6 dB offset addition process.

If the log energy exceeds this threshold - which would correspond to a frame of speech rather than background noise - the current background noise estimate is not updated, and the background noise updating process terminates. If, however, the log energy does not exceed the noise threshold level - which would correspond to a detected minima in the post-processed signal indicating that only noise is present - the background noise spectral estimate is updated in block 745. This corresponds to the closing of channel switch 575 in response to a positive valley detect signal from energy valley detector 570. This updating process consists of providing a time-averaged value of the pre-processed channel energy estimate for the particular channel by smoothing the estimate (in smoothing filter 580), and storing these time averaged values as per-channel noise estimates (in energy estimate storage register 585). The operation of background noise estimator block 621 ends for the particular frame being processed by proceeding to block 629 and 630 to obtain a new frame.

Now, to describe the operation and structure of the present invention in detail (and referring again to FIG. 1), improved noise suppression 800 incorporates changes to the aforementioned Borth noise suppression system in three basic area: (a) the updating of background noise estimates by voice metric calculator 810; (b) the modification of SNR estimates by channel SNR modifier 820; and (c) utilization of SNR threshold block 830 to offset the gain rise of each channel. Each of these improvements will be described in terms of the block diagram of Figure 1, and in the terms of the flowchart of Figure 4a - 4f.

Voice metric calculator 810 replaces the valley detector circuitry of the previous system. A voice metric is essentially a measurement of the overall voice-like characteristics of all the channel energies. In the preferred embodiment, voice metric calculator 810 is implemented as a look-up table which translates the individual channel SNR estimates at 235 into voice metric values. The voice metric values are used internally to determine when to update the background noise estimate, by closing channel switch 575 for one frame. As used herein, updating the background noise estimate is defined as partially modifying the old background noise estimate with a new estimate using, for example, a 10%/90% new-to-old estimate ratio. The voice metric values are also used in the channel SNR modifying process as will subsequently be described.

From the perspective of making a background noise update decision, a frame having high energy, which is typically indicative of a speech frame, could also mean that a narrowband noise transient or a sudden increase in the background noise level has occurred. Therefore, the present invention characterizes the frame energy as a voice metric sum, VMSUM, and utilizes this multi-channel energy parameter to perform the updating decision. The process utilizes a voice metric table which may be represented as a curve as shown in figure 2.

Figure 2 is a graph illustrating the characteristic curve of the voice metrics for a particular channel. the horizontal axis represents SNR estimate indices. Each SNR estimate index value represents three-eighths (3/8) dB signal-to noise ratio. Hence, an SNR estimate index of 10 represents 3.75 dB SNR. The vertical axis represents voice metric values VM(CC) for each of the N Channels. Note that a voice metric of 2 is produced for an SNR index of 1. Also note that the curve is not linear, since a channel energy has more voice-like characteristics at higher SNR's.

First, the raw SNR estimates are used to index into the voice metric table to obtain a voice metric value VM(CC) for each channel. Second, the individual channel voice metric values are summed to create the total of all individual channel voice metric values, called the voice metric sum VMSUM. Third, VMSUM is compared to an UPDATE THRESHOLD representative of a voice metric total that is deemed to be noise. If the multi-channel energy parameter VMSUM is less that the UPDATE THRESHOLD, the particular frame has very few voice-like characteristics, and is most probably noise. Therefore, a background noise update is performed by closing channel switch 575 for the particular frame. The most recent voice metric sum VMSUM is also made available to channel SNR modifier 820 via line 815 for use in the modification algorithm.

In the preferred embodiment, the UPDATE THRESHOLD is set to a total voice metric sum value of 32. Since the minimum value in the voice metric table is 2, the minimum sum for 14 channels is 28. The voice metric tables values remain at 2 until an SNR index of 12 (or 4.5 dB SNR) is reached. This means that an increased level of broadband

noise (individual channels each having SNR values not greater that 4.125 dB) will still generate a sum of 28. Since the UPDATE-THRESHOLD of 32 would not then be exceeded, the broadband noise voice metric will be correctly classified as noise, and a background noise update will be performed. Conversely, any single channel having an SNR index value greater than 24 (or at least 9.0 dB SNR) would cause the VMSUM to exceed the UPDATE THRESHOLD, and result in a voice or narrowband noise burst decision.

Many variations of the voice metric table are possible, as different types of metrics may be compensated for by the proper selection of the UPDATE THRESHOLD. Furthermore, the sensitivity of the speech/noise decision may also be chosen for a particular application. For example, in the preferred embodiment, the threshold may be adjusted to accommodate any single channel having an SNR value as sensitive as 4.5 dB to as insensitive as 15 dB. The corresponding UPDATE THRESHOLD would then be set within the range of 29 to 41.

In addition to performing the speech/noise decision utilizing voice metrics, voice metric calculator 810 keeps track of the time that has expired since the last background noise update. An update counter is tested on each frame to see if more than a given number of names, each representing a predetermined time, has passed since the previous update. In the preferred embodiment utilizing 10 millisecond frames, if the update counter reaches 100 --corresponding to a timing threshold of 1 second without updates-- an update is performed regardless of the voice metric decision. However, any timing threshold within the range of 0.5 second to 4 seconds would be practical. As previously mentioned this timing parameter test is used to prevent any sudden, large increases in noise level from being indefinitely interpreted as voice.

The basic function of channel SNR modifier 820 is to eliminate the detrimental effects of narrowband noise bursts on the noise suppression system. A narrowband noise burst may be defined as a momentary increase in channel energy for only a few channels. In the preferred embodiment, a high energy level above a 6 dB SNR threshold in fewer than 5 of the upper 10 channels is classified as a narrowband noise burst. Such a noise burst would normally create high gain values for only a few number of channels, which results in the "running water" type of background noise flutter described above.

Raw SNR estimates at 235 are applied to the input of channel SNR modifier 820, and modified SNR estimates are output at 825. Basically, SNR modifier 820 counts the number of channels which have channel SNR index values which exceed an index threshold. In the preferred embodiment, the index threshold is set to correspondto an SNR value within the range of 4 dB to 10 dB, preferably 6 dB SNR. If the number of channels is below a predetermined count threshold, then the decision to modify the SNR's is made. The count threshold represents a relatively few number of channels, i.e., not greater than 40% of the total number of channels N. In the preferred embodiment, the count threshold is set to 5 of the 10 measured channels. During the modification process itself, channel SNR modifier 820 either reduces the SNR of only those particular channels having an SNR index less than a SETBACK THRESHOLD (indicative of a narrowband noise channel), or reduces the SNR of all the channels if the voice metric sum is less than a metric threshold (indicative of a very weak energy frame). Hence, the channels containing the narrowband noise burst are attenuated so as to prevent them from detrimentally affecting the gain table look-up function.

SNR threshold block 830 proves a predetermined SNR threshold for each channel which must be exceeded by the modified channel SNR estimates before a high gain value can be produced. Only SNR estimates which have a value above the SNR threshold are directly applied to the gain table sets. Therefore, small background noise fluctuation are not allowed to produce gain values which represent voice. This implementation of an SNR threshold essentially presents an offset in the gain rise for channels having low signal-to-noise ratio. Preferably, the SNR threshold would be set within the range of 1.5 dB to 5 dB SNR to eliminate minor noise fluctuations. The SNR threshold may be implemented as a separate element as shown in Figure 1, or it may be implemented as a "dead zone" in the characteristic gain curve for each gain table set 590.

Figure 3 graphically illustrates the function of SNR threshold block 830, as well as the attenuation function of the channel gain values in each gain table set. On the horizontal axis, modified SNR estimates are shown in dB as would be output from channel SNR modifier 820 at 825. The vertical axis represents the channel gain (attenuation) as would be observed at the output of channel gain modifier 250 at 255. A maximum amount of background noise attenuation is achieved for channels having a minimum gain value. Note that SNR threshold block 830 is shown as "dead zone" or offset in the gain rise curve of approximately 2.25 dB. Hence, an SNR estimate must exceed this threshold before the channel gain can rise above the minimum gain level shown. Also note that two curves are illustrated, each having a different minimum gain level. Upper curve labeled A group represents a low channel group, e.g., consisting of channels 1-4 in the preferred embodiment, while group B represents the higher frequency channels 5-14.

As evident from the graph, the low frequency channels have a minimum gain value of-13.1 dB, while the upper frequency channels have a minimum gain value of -20.7 dB. It has been found that less voice quality degradation occurs when the channels are divided into such groups. Although only two different gain curves are used in the preferred embodiment for gain table set number 1, it may prove advantageous to provide each channel with a different characteristic gain curve. Furthermore, as explained in the referenced Borth patent, multiple gain table sets are used to allow a wider choice of channel gain values depending on the particular background noise environment. Noise level quantizer

555 utilizes hysteresis to select a particular gain table set based upon the overall background noise estimates. The gain table selection signal, output from noise level quantizer 555, is applied to gain table switch 595 to implement the gain table selection process. Accordingly, one of a plurality of gain table sets 590 may be chosen as a function of overall average background noise level.

These noise suppression improvements eliminate the variability of the background noise suppression without requiring a large amount of gain smoothing. Background noise attenuation within the range of 10 dB to 25 dB is readily achieved with the present invention. With the improvements, the system requires gain smoothing having a time constant of only 10 to 20 milliseconds to obtain a flat or "white" residual noise background. Previous techniques required 40 to 60 millisecond time constant gain smoothing, which not only resulted in imperfect flutter reduction, but also substantially degraded the voice quality.

Since the overall operation of the improved noise suppression system is similar to that described in the previous descriptiuon of the prior art US-A-4,628,529, the generalized flow diagram illustrated in figures 5a/b of this application are again used to describe the present invention. The general organization of the operation of the present invention may still be organized in three functional groups: noise suppression loop -- sequence block 604 of previously described prior art figure 5a, which is described in greater detail in prior art Figure 6a of the present application; automatic gain selector -- sequence 615 of previously described prior art figure 5b, which has been modified for the present invention; and automatic background noise estimator -- sequence 621 of prior art Figure 5b, which has also been modified in the present invention. The detailed flowcharts of Figure 4a through 4f of the present application may be substituted for sequence blocks 615 and 621 of prior art Figure 5b to describe the operation of improved noise suppression system 800. Hence, prior art Figure 5a and 6a of the previously described Borth patent (4,628,529) describe the noise suppression loop performed on a sample-by-sample basis, while Figures 4a through 4f of the present invention describe the channel gain selection process and the background noise estimate update process performed on a frame-by-frame basis.

Referring now to Figure 4a, the operation of improved noise suppression system 800 begins from the "YES" output of decision step 614 of the aforementioned Figure 6a. Hence the actual spectral gain modification function for the particular frame has already been performed on a sample-by-sample basis utilizing gain values from the previous frame. Sequence 850 serves to generate the SNR estimates available at 235. First of all, the channel count CC is set equal to 1 in step 851. Next, the voice metric sum variable VMSUM is initialized to zero in step 852. Step 853 calculates the raw signal-to-noise ratio SNR for the particular channel as an SNR estimate index value INDEX(CC). The SNR calculation is simply a division of the per-channel energy estimates (signal-plus-noise) available at 225, by the per-channel background noise estimates (noise) at 325. However, other estimates of the signal-to-noise threshold may alternatively be used. Therefore, step 853 simply divides the current stored channel energy estimate (obtained from flowchart step 707 of the aforementioned Figure 7a) by the current background noise estimate BNE(CC) from the previous frame.

In sequence 860, the voice metrics are calculated. First, the voice metric table for the particular channel is indexed in step 861 using the raw SNR estimate index INDEX(CC). The voice metric table is read in step 862 to obtain a voice metric value VM(CC) for that particular channel. This individual channel voice metric value is added to the voice metric sum VMSUM in step 863. The channel count CC is incremented in step 864, and tested in step 865. If the voice metrics for all N channels have not been calculated, control returns to step 853.

Sequence 870 illustrates the background noise estimate update decision process performed by voice metric calculator 810. The voice metric sum VMSUM is compared to UPDATE THRESHOLD in step 871. If VMSUM is less than or equal to UPDATE THRESHOLD, then the frame is probably a noise frame. TIMER FLAG is reset in step 872, and the update counter UC is reset in step 873. Control proceeds to 878 where the UPDATE FLAG is set true, which means that a background noise estimate update will be performed for the current frame.

If VMSUM is greater that the UPDATE THRESHOLD, the frame is probably a voice frame. Nevertheless, step 874 tests the TIMBER FLAG to see if a sudden, loud increase in background noise has been interpreted as speech. If the TIMER FLAG is true, the one second time interval was exceeded a number of frames ago, and background noise estimate updating is still required. This is due to the fact that only a partial background noise update is performed for each frame. If the TIMER FLAG is not true, the update counter UC is incremented in step 875, and tested in step 876. If 100 frames have occurred since the last background noise estimate update, the TIMER FLAG is set true in step 877, and the BNE UPDATE FLAG is set true in step 878. A series of partial background noise estimate updates are then performed until the voice metric sum VMSUM again falls below the UPDATE THRESHOLD. Note that the only place in the flowchart that the TIMBER FLAG is reset is in step 872, when the voice metric sum VMSUM again resembles noise. If the update counter UC has not reached 100 frames, the instant frame is deemed to be a voice frame, and no background noise update is performed.

Referring now to sequence 880 of Figures 4b and 4c, the decision to modify the channel signal-to-noise ratios is performed next. An index counter variable IC is initialized in step 881. The channel counter CC is set equal to 5 in step 882, so as to count only the upper 10 of the 14 channels having a high energy. The raw SNR estimate index INDEX

(CC) is tested in step 883 to see if it has reached an INDEX THRESHOLD which would correspond to approximately 6 dB SNR. Here, the assumption is made that at least 5 of the upper 10 channels of a voice frame should contain energy having an SNR of at least 6 dB. If the particular channel SNR INDEX(CC) is above the INDEX THRESHOLD, the index count IC is incremented in step 884. If not, the channel count CC is incremented in step 885 and tested in step 886 to look at the next channel.

When all 10 upper channels have been measured, index count IC represents the number of channels having an SNR estimate index higher than the INDEX THRESHOLD. The index count IC is then tested against a COUNT THRESHOLD in step 887. If IC indicates that more channels than the COUNT THRESHOLD, e.g., 5 of the upper 10 channels, contain sufficient energy, then the frame is probably a voice frame, and the MODIFY FLAG is set false in step 889 to prevent channel SNR modification. If only a few channels contain high energy, which would be representative of a frame of narrowband noise, then the MODIFY FLAG is set true in step 888.

Sequence 890 describes the SNR modification process performed by channel SNR modifier block 820. Initially, the MODIFY FLAG is tested in step 891. If it is false, the channel SNR modification process is by-passed. If the MODIFY FLAG is true, the channel counter CC is initialized in step 892. Next, each channel SNR estimate index is tested in step 893 to see if it is less than or equal to a SETBACK THRESHOLD. The SETBACK THRESHOLD, which may have a value corresponding to 6 dB SNR, represents the maximum SNR estimate which is representative of background noise flutter. Only channels having low SNR estimate index pass this test. However, even if the channel index is greater than the SETBACK THRESHOLD, the voice metric sum VMSUM is again tested in step 894. If VMSUM is less than or equal to a METRIC THRESHOLD, which corresponds to a representative total voice metric of a narrowband noise frame, the INDEX (CC) is modified in step 895 by setting it equal to the minimum index value of 1. The channel counter CC is incremented in step 896 and tested in step 897 to see if all the channels have been tested. If not, control returns to step 893 to test the next channel index. Hence, a frame containing either channel energy fluctuations or narrowband noise is modified such that the frame does not produce undesirable gain variations.

Sequence 900 performs the function of SNR threshold block 830. The channel counter CC is initialized in step 901. The SNR index for the particular channel is tested against an SNR THRESHOLD in step 902. In the preferred embodiment, the SNR THRESHOLD represents an index value corresponding to 2.25 dB SNR. If INDEX (CC) is above the SNR THRESHOLD, it may be used to index the gain table. If not, the index value is again set equal to 1 in step 903, which represents the minimum index value. The channel counter CC is incremented in step 904 and tested in step 905. This SNR threshold testing process serves to reduce minor background noise variations in all the channels.

Referring now to sequence 910 of Figure 4d, the gain table sets are chosen by noise level quantizer 555 and gain table switch 595. In step 911, the channel counter CC is initialized, and in step 912, a variable called background noise estimate sum, BNESUM, is intialized. In step 913, the current background noise estimate BNE(CC) is obtained for each channel, and added to BNESUM in step 914. Step 915 increments the channel counter CC, and step 916 tests the channel counter to see if the background noise estimates for all N channels have been totaled.

In step 917, BNESUM is compared to a first background noise estimate threshold. If it is greater than BNE THRESHOLD 1, then gain table set number 1 is selected in step 918. Similarly, step 919 again tests BNESUM to see if it is greater than the lower value of BNE THRESHOLD 2. If BNESUM is greater than BNE THRESHOLD 2 but less than BNE THRESHOLD 1, then gain table set number 2 is selected in step 920. Otherwise, gain table set number 3 is selected in step 921. Hence, gain table sets 590 are selected as a function of overall average background noise level.

Sequence 930 describes the steps for obtaining raw gain values RG (CC) from the gain table sets 590. Step 931 sets the channel counter CC equal to 1. The selected gain table is indexed in steo 932 using the channel SNR estimate index INDEX(CC) which has passed the SNR modification and threshold tests. The raw gain value RG(CC) is obtained from the selected gain table in step 933, and is then stored in step 934 for use as the gain values for the next frame of noise suppression. The channel counter CC is incremented in step 935, and tested in step 936 as before. As described in U.S. Patent No. 4,630,305, the raw gain values for each channel at 535 are then applied to gain smoothing filter 530 for smoothing on a per-sample basis.

Finally, sequence 940 describes the actual background noise estimate updating process performed in block 420 of Figure 1. Step 941 initially tests the UPDATE FLAG to see if a background noise estimate should be performed. If the UPDATE FLAG is false, then the frame is a voice frame and no background noise update can occur. Otherwise, the background noise update is performed - which is simulated by closing channel switch 575 -- during a noise frame. In step 942, the UPDATE FLAG is reset to false. Steps 942 through 945 serve to update the current background noise estimate in each of the N channels via the equation:

$$E(i,k) = E(u,k-1) + SF[(PE(i)-E(i,k-1)],$$

$$I=1,2,\ldots, N$$

where E(i,k) is the current energy noise estimate for channel (i) at time (k), E(i, k-l) is the old energy noise estimate

for channel (i) at time (k-l), PE(i) is the current pre-processed energy estimate for channel (i), and SF is the smoothing factor time constant used in smoothing the background noise estimates. Therefore, E(i, k-l) is stored in energy estimate storage register 585, and the SF term performs the function of smoothing filter 580. In the present embodiment, SF is selected to be 0.1 for a 10 millisecond frame duration.

Step 943 initializes the channel count CC to 1. Step 944 performs the above equation in terms of the current background noise estimate available at 325, the old background noise estimate available at 325, the old background noise estimate OLD BNE (CC) stored in energy estimate storage register 585, and the new background noise estimate NEW BNE(CC) available from switch 575. Step 945 increments the channel counter CC, and step 946 tests to see of all N channels have been processed. If true, the background noise estimate update is completed, and operation is returned to step 629 of Figure 6b of the aforementioned Borth patent to reset the sample counter and increment the frame counter. Control then returns to perform noise suppression on a sample-by-sample basis for the next frame.

In review, it can now be seen that the present invention provides the following improvements: (a) a reduction in background noise flutter by offsetting the gain rise of the gain tables until a certain SNR value is obtained; (b) immunity to narrowband noise bursts through modification of the SNR estimates based on the voice metric calculation and the channel energies; and (c) more accurate background noise estimates via performing the update decision based on the overall voice metric and the time interval since the last update.

While specific embodiments of the present invention have been shown and described herein, further modification and improvements may be made by those skilled in the art. For example, the operational flow is described herein as performed in real time. However, due to inherent hardware limitations, previous background noise estimates for channel gain values may be stored for use in the next frame. All such modification which retain the basic underlying principles disclosed and claims herein are within the scope of this invention.

Claims

1. A noise suppression system (800) for attenuating the background noise from a noisy input signal (205) to produce a noise-suppressed output signal (265), said noise suppression system comprising:

   means (210) for separating the input signal (205) into a plurality of pre-processed signals representative of selected frequency channels;
   means (220) for generating estimates of the signal-plus-noise energy and the noise energy in each individual channel; and
   means (310, 810) for producing a gain value for each individual channel in response to said channel energy estimates;

   the system (800) characterized in that:

   said gain values have a minimum gain value for each channel;
   said means (310, 810) for producing a gain value includes threshold means for allowing gain values above said minimum gain value to be produced only when said signal-plus-noise energy estimates exceed said noise energy estimates by a predetermined amount; and
   means for modifying the gain of each of said plurality of pre-processed signals in response to said gain values to provide a plurality of post-processed signals.

2. The noise suppression system according to claim 1, wherein said means (310, 810) for producing a gain value produces gain values based upon the signal-to-noise ratio (SNR) of said channel energy estimates, and wherein said SNR estimates are compared with a predefined SNR threshold such that channels having SNR estimates below said SNR threshold produce minimum gain values.

3. The noise suppression system according to claim 2, wherein said predefined SNR threshold corresponds to an SNR value within the range of 1.5 dB to 5 dB SNR.

4. The noise suppression system according to claim 3, wherein said predefined SNR threshold corresponds to an SNR value of approximately 2.25 dB SNR.

5. The noise suppression system according to any preceding claim, wherein said gain modifying means provides a maximum amount of attenuation of the pre-proposed signal in a particular channel having a minimum gain value.

6. The noise suppression system according to any preceding claim, wherein gain values produce a higher amount of attenuation for high frequency channels than low frequency channels.

7. The noise suppression system according to any preceding claim 1, wherein said means (310, 810) for producing a gain value further includes a plurality of gain tables, each gain table having predetermined individual channel gain values corresponding to said individual channel energy estimates, and gain table selection means for automatically selecting one of said plurality of gain tables as a function of the overall average background noise level of said input signal.

8. The noise suppression system according to any preceding claim, further including means for combining said plurality of post-processed signals to produce said noise-suppressed output signal.

9. The noise suppression system of any preceding claim, wherein said means (220) for gemating estimates includes:

means for generating and storing an estimate of the background noise power spectral density of said pre-processed signals, said background noise estimate generating means including means for modifying said background noise estimate in response to a timing parameter indicative of the time interval since the previous background noise estimate modification; and
means for generating an estimate of the signal-to-noise ration (SNR) in each individual channel based upon said modified background noise estimates;
wherein said means (310, 810) for producing a gain value for each individual channel produces each gain value in response to said channel SNR estimates.

10. The noise suppression system according to claim 9, wherein said background noise estimate modifying means includes means for producing said timing parameter, and means for comparing said timing parameter to a predetermined timing threshold such that a background noise estimates modification is performed when said timing parameter exceeds said timing threshold.

11. The noise suppression system according to claim 10, wherein said background noise estimate modifying means further includes means for generating an estimate of the energy in each individual channel, and means for producing a multi-channel energy parameter in response to the total value of all individual channel energy estimates.

12. The noise suppression system according to claim 11, wherein said multi-channel energy parameter producing means accommodates for minor variations in individual channel energy estimates such that said minor variations do not significantly affect said multi-channel energy parameter.

13. The noise suppression system according to claim 11 or 12, wherein said background noise estimates modifying means further includes means for comparing said multi-channel energy parameter to a predetermined energy threshold such that a background noise estimates modification is performed when said multi-channel energy parameter is less than said energy threshold.

14. The noise suppression system according to claim 13, wherein said background noise estimates modifying means modifies said background noise estimates in response to said timing parameter regardless of said multi-channel energy parameter.

15. The noise suppression system according to any one of claims 11 to 14, wherein said multi-channel energy parameter is generated by translating said individual channel SNR estimates the individual channel voice metrics, the voice metric sum being a measurement of the overall voice-like characteristics of the energy in all channels.

16. The noise suppression system according to any preceding claim, further comprising:

means (815, 820, 830) for monitoring said channels energy estimates and for distinguishing narrowband noise bursts from speech energy and background noise energy, thereby producing a modification signal (835);
means (590) for selectively modifying said channels energy estimates in response to said modification signal (835) such that channel energy estimates representative of narrowband noise bursts are modified;
means for producing a gain value for each individual channel in response to each modified channel energy estimate; and
means for modifying the gain of each of said plurality of pre-processed signals in response to said gain values

to provide a plurality of post-processed signals.

17. The noise suppression system according to claim 16, wherein said modification signal is indicative of the total number of individual channels having energy estimates exceeding a predetermined energy threshold.

18. The noise suppression system according to claim 16 or 17, wherein said channel energy estimates modifying means includes means (830) for comparing said modification signal to a predetermined count threshold such that a channel energy estimates modification is performed when said total number of individual channels is less than said count threshold.

19. The noise suppression system according to claim 16, 17 or 18, wherein said gain modifying means provides a maximum amount of attenuation of the pre-processed signal in a particular channel having a modified channel energy estimate.

20. A method of attenuating the background noise from a noisy input signal (205) to produce a noise-suppressed output signal (265) in a noise suppression system (800) comprising the steps of:

separating (850) the input signal into a plurality of pre-processed signals representative of a number N of selected frequency channels;
generating an estimate of the energy in each individual channel (853);
generating and storing an estimate of the background noise power spectral density of said pre-processed signals; and
generating an estimate of the signal-to-noise ratio (SNR) in each individual channel based upon said background noise estimates and said channel energy estimates;

the method characterized by the steps of:

producing (861, 862) a gain value for each individual channel in response to said channel SNR estimates, wherein said gain values have a range of minimal values;
and wherein said gain values producing step includes the steps of:
providing a predefined SNR threshold and comparing (902) said channel SNR estimates to said predefined SNR threshold such that channels having SNR estimates below said SNR threshold produce gain values within said minimal range; and
modifying (910) the gain of each of said plurality of pre-processed signals in response to said gain values to provide a plurality of post-processed signals.

Patentansprüche

1. Rauschunterdrückungssystem (800) zur Dämpfung des Hintergrundrauschens von einem verrauschten Eingangssignal (205), um ein rauschunterdrücktes Ausgangssignal (265) zu erzeugen, wobei dieses Rauschunterdrükkungssystem umfaßt:

- Mittel (210) zur Aufteilung des Eingangssignals (205) in eine Vielzahl von vorverarbeiteten Signalen, die ausgewählte Frequenzkanäle darstellen;
- Mittel (220) zur Erzeugung von Abschätzungen der Signalplus-Rausch-Energie und der Rauschenergie in jedem einzelnen Kanal; und
- Mittel (310, 810) zur Erzeugung eines Verstärkungswerts für jeden einzelnen Kanal in Reaktion auf diese Kanalenergieabschätzungen;

wobei das System (800) dadurch gekennzeichnet ist, daß:

- diese Verstärkungswerte für jeden Kanal einen minimalen Verstärkungswert haben;
- diese Mittel (310, 810) zur Erzeugung eines Verstärkungswerts Schwellenwertmittel enthalten, die gestatten, daß Verstärkungswerte oberhalb dieses minimalen Verstärkungswerts nur erzeugt werden, wenn diese Signalplus-Rausch-Energieabschätzungen diese Rauschenergieabschätzungen um einen vorbestimmten Betrag überschreiten; und
- Mittel zur Modifizierung der Verstärkung jedes aus der Vielzahl vorverarbeiteter Signale in Reaktion auf diese

Verstärkungswerte, um eine Vielzahl nachverarbeiteter Signale bereitzustellen.

2. Rauschunterdrückungssystem nach Anspruch 1, wobei diese Mittel (310, 810) zur Erzeugung eines Verstärkungswerts auf der Grundlage des Signal-Rausch-Abstands (SNR) dieser Kanalenergieabschätzungen Verstärkungswerte erzeugt und wobei diese SNR-Abschätzungen mit einem vordefinierten SNR-Schwellenwert verglichen werden, so daß Kanäle, die SNR-Abschätzungen unter diesem SNR-Schwellenwert haben, minimale Verstärkungswerte erzeugen.

3. Rauschunterdrückungssystem nach Anspruch 2, wobei dieser vordefinierte SNR-Schwellenwert einem SNR-Wert innerhalb des Bereichs von 1,5 dB bis 5 dB SNR entspricht.

4. Rauschunterdrückungssystem nach Anspruch 3, wobei dieser vordefinierte SNR-Schwellenwert einem SNR-Wert von ungefähr 2,25 dB SNR entspricht.

5. Rauschunterdrückungssystem nach einem der vorhergehenden Ansprüche, wobei diese verstärkungsmodifizierenden Mittel einen maximalen Betrag der Dämpfung des vorverarbeiteten Signals in einem bestimmten Kanal, der einen minimalen Verstärkungswert hat, gewährleisten.

6. Rauschunterdrückungssystem nach einem der vorhergehenden Ansprüche, wobei die Verstärkungswerte einen größeren Betrag der Dämpfung für Kanäle mit hohen Frequenzen als für Kanäle mit niedrigen Frequenzen erzeugen.

7. Rauschunterdrückungssystem nach einem der vorhergehenden Ansprüche, wobei diese Mittel (310, 810) zur Erzeugung eines Verstärkungswerts weiter eine Vielzahl von Verstärkungstabellen enthalten, wobei jede Verstärkungstabelle vorbestimmte einzelne Kanalverstärkungswerte entsprechend dieser einzelnen Kanalenergieabschätzungen hat, und wobei Verstärkungstabellenauswahlmittel zur automatischen Auswahl einer Verstärkungstabelle aus dieser Vielzahl von Verstärkungstabellen als eine Funktion des gesamtdurchschnittlichen Hintergrundrauschpegels dieses Eingangssignals vorhanden sind.

8. Rauschunterdrückungssystem nach einem der vorhergehenden Ansprüche, das weiter Mittel zur Kombinierung dieser Vielzahl nachverarbeiteter Signale enthält, um dieses rauschunterdrückte Ausgangssignal zu erzeugen.

9. Rauschunterdrückungssystem nach einem der vorhergehenden Ansprüche, wobei diese Mittel (220) zur Erzeugung der Abschätzungen enthalten:

- Mittel zur Erzeugung und Speicherung einer Abschätzung der Hintergrundrauschleistungsspektraldichte dieser vorverarbeiteten Signale, wobei diese hintergrundrauschabschätzungserzeugenden Mittel Mittel zur Modifizierung dieser Hintergrundrauschabschätzung in Reaktion auf einen Zeitparameter enthalten, der eine Anzeige des Zeitintervalls seit der vorherigen Hintergrundrauschabschätzungsmodifikation ist; und
- Mittel zur Erzeugung einer Abschätzung des Signal-Rausch-Abstands (SNR) in jedem einzelnen Kanal auf der Grundlage dieser modifizierten Hintergrundrauschabschätzungen;

wobei diese Mittel (310, 810) zur Erzeugung eines Verstärkungswerts für jeden einzelnen Kanal jeden Verstärkungswert in(Reaktion auf diese Kanal-SNR-Abschätzungen erzeugen.

10. Rauschunterdrückungssystem nach Anspruch 9, wobei diese hintergrundrauschabschätzungsmodifizierenden Mittel Mittel zur Erzeugung dieses Zeitparameters und Mittel zum Vergleich dieses Zeitparameters mit einem vorbestimmten Zeitgeberschwellenwert enthalten, so daß eine Hintergrundrauschabschätzungsmodifikation durchgeführt wird, wenn dieser Zeitparameter diesen Zeitgeberschwellenwert überschreitet.

11. Rauschunterdrückungssystem nach Anspruch 10, wobei diese hintergrundrauschabschätzungsmodifizierenden Mittel weiter Mittel zur Erzeugung einer Abschätzung der Energie in jedem einzelnen Kanal und Mittel zur Erzeugung eines Mehrkanalenergieparameters in Reaktion auf den Gesamtwert aller einzelnen Kanalenergieabschätzungen enthalten.

12. Rauschunterdrückungssystem nach Anspruch 11, wobei diese mehrkanalenergieparametererzeugenden Mittel sich geringen Veränderungen der einzelnen Kanalenergieabschätzungen anpassen, so daß diese geringen Veränderungen diesen Mehrkanalenergieparameter nicht wesentlich beeinflussen.

13. Rauschunterdrückungssystem nach Anspruch 11 oder 12, wobei diese hintergrundrauschabschätzungsmodifizierenden Mittel weiter Mittel zum Vergleich dieses Mehrkanalenergieparameters mit einem vorbestimmten Energieschwellenwert enthalten, so daß eine Hintergrundrauschabschätzungsmodifikation durchgeführt wird, wenn dieser Mehrkanalenergieparameter kleiner als dieser Energieschwellenwert ist.

14. Rauschunterdrückungssystem nach Anspruch 13, wobei diese hintergrundrauschabschätzungsmodifizierenden Mittel diese Hintergrundrauschabschätzungen in Reaktion auf diesen Zeitparameter modifizieren, ungeachtet dieses Mehrkanalenergieparameters.

15. Rauschunterdrückungssystem nach einem der Ansprüche 11 bis 14, wobei dieser Mehrkanalenergieparameter dadurch erzeugt wird, daß diese einzelnen Kanal-SNR-Abschätzungen die einzelnen Kanalsprachmaße übersetzen, wobei die Sprachmaßsumme eine Messung der gesamten sprachähnlichen Charakteristiken der Energie in allen Kanälen ist.

16. Rauschunterdrückungssystem nach einem der vorhergehenden Ansprüche, weiter umfassend:

- Mittel (815, 820, 830) zur Überwachung dieser Kanalenergieabschätzungen und zur Unterscheidung der Schmalbandrauschbündel von Sprachenergie und Hintergrundrauschenergie, wodurch ein Modifikationssignal (835) erzeugt wird;
- Mittel (590) zur selektiven Modifizierung dieser Kanalenergieabschätzungen in Reaktion auf dieses Modifikationssignal (835), so daß Kanalenergieabschätzungen, die Schmalbandrauschbündel darstellen, modifiziert werden;
- Mittel zur Erzeugung eines Verstärkungswerts für jeden einzelnen Kanal in Reaktion auf jede modifizierte Kanalenergieabschätzung; und
- Mittel zur Modifizierung der Verstärkung jedes aus dieser Vielzahl vorverarbeiteter Signale in Reaktion auf diese Verstärkungswerte, um eine Vielzahl nachverarbeiteter Signale bereitzustellen.

17. Rauschunterdrückungssystem nach Anspruch 16, wobei dieses Modifikationssignal eine Anzeige der Gesamtanzahl der einzelnen Kanäle ist, die Energieabschätzungen haben, die einen vorbestimmten Energieschwellenwert überschreiten.

18. Rauschunterdrückungssystem nach Anspruch 16 oder 17, wobei diese kanalenergieabschätzungsmodifizierenden Mittel Mittel (830) zum Vergleich dieses Modifikationssignals mit einem vorbestimmten Zählerschwellenwert enthalten, so daß eine Kanalenergieabschätzungsmodifikation durchgeführt wird, wenn diese Gesamtanzahl der einzelnen Kanäle kleiner als dieser Zählerschwellenwert ist.

19. Rauschunterdrückungssystem nach Anspruch 16, 17 oder 18, wobei diese verstärkungsmodifizierenden Mittel einen maximalen Betrag der Dämpfung des vorverarbeiteten Signals in einem bestimmten Kanal, der eine modifizierte Kanalenergieabschätzung hat, gewährleisten.

20. Verfahren zur Dämpfung des Hintergrundrauschens eines verrauschten Eingangssignals (205), um in einem Rauschunterdrückungssystem (800) ein rauschunterdrücktes Ausgangssignal (265) zu erzeugen, wobei das Verfahren die folgenden Schritte umfaßt:

- Aufteilung (850) des Eingangssignals in eine Vielzahl vorverarbeiteter Signale, die durch eine Anzahl von N ausgewählten Frequenzkanälen dargestellt werden;
- Erzeugung einer Abschätzung der Energie in jedem einzelnen Kanal (853);
- Erzeugung und Speicherung einer Abschätzung der Hintergrundrauschleistungsspektraldichte dieser vorverarbeiteten Signale; und
- Erzeugung einer Abschätzung des Signal-Rausch-Abstands (SNR) in jedem einzelnen Kanal auf der Grundlage dieser Hintergrundrauschabschätzungen und dieser Kanalenergieabschätzungen;

wobei das Verfahren durch die folgenden Schritte gekennzeichnet ist:

- Erzeugung (861, 862) eines Verstärkungswerts für jeden einzelnen Kanal in Reaktion auf diese Kanal-SNR-Abschätzungen, wobei diese Verstärkungswerte einen Bereich minimaler Werte haben;

und wobei dieser verstärkungswerterzeugende Schritt die folgenden Schritte enthält:

- Bereitstellung eines vordefinierten SNR-Schwellenwerts und Vergleich (902) dieser Kanal-SNR-Abschätzungen mit diesem vordefinierten SNR-Schwellenwert, so daß Kanäle, die SNR-Abschätzungen unter diesem SNR-Schwellenwert haben, Verstärkungswerte innerhalb dieses minimalen Bereichs erzeugen; und
- Modifizierung (910) der Verstärkung jedes aus dieser Vielzahl vorverarbeiteter Signale in Reaktion auf diese Verstärkungswerte, um eine Vielzahl nachverarbeiteter Signale bereitzustellen.

Revendications

1. Système d'élimination de bruit (800) permettant d'atténuer le bruit de fond provenant d'un signal d'entrée bruyant (205) pour produire un signal de sortie dont le bruit a été supprimé (265), ledit système d'élimination de bruit comprenant :

   un moyen (210) permettant de séparer le signal d'entrée (205) en une pluralité de signaux prétraités représentant les voies de fréquence choisies ;
   un moyen (220) permettant de générer des estimations de l'énergie du signal plus bruit et de l'énergie du bruit dans chaque voie individuelle ; et
   un moyen (310, 810) permettant de produire une valeur de gain pour chaque voie individuelle en réponse auxdites estimations d'énergie de voie ;
   le système (800) étant caractérisé en ce que : lesdites valeurs de gain présentent une valeur de gain minimum pour chaque voie ;
   ledit moyen (310, 810) permettant de produire une valeur de gain comporte un moyen formant seuil permettant à des valeurs de gain supérieures à ladite valeur de gain minimum d'être produites uniquement lorsque lesdites estimations d'énergie de signal-plus-bruit dépassent lesdites estimations d'énergie de bruit d'une quantité prédéterminée ; et
   un moyen permettant de modifier le gain de chaque signal prétraité de ladite pluralité de signaux prétraités en réponse auxdites valeurs de gain afin de fournir une pluralité de signaux post-traités.

2. Système d'élimination de bruit selon la revendication 1, dans lequel ledit moyen (310, 810) permettant de produire une valeur de gain produit des valeurs de gain qui sont fonction du rapport signal/bruit (SNR) desdites estimations d'énergie de voie, et dans lequel lesdites estimations de SNR sont comparées à un seuil de SNR prédéfini de sorte que les voies présentant des estimations de SNR inférieures audit seuil de SNR produisent des valeurs de gain minimum.

3. Système d'élimination de bruit selon la revendication 2, dans lequel ledit seuil de SNR prédéfini correspond à une valeur de SNR comprise entre 1,5 dB et 5 dB.

4. Système d'élimination de bruit selon la revendication 3, dans lequel ledit seuil de SNR prédéfini correspond à une valeur de SNR d'environ 2,25 dB.

5. Système d'élimination de bruit selon l'une quelconque des revendications précédentes, dans lequel ledit moyen de modification de gain fournit une quantité maximum d'atténuation du signal pré-proposé dans une voie particulière ayant une valeur de gain minimum.

6. Système d'élimination de bruit selon l'une quelconque des revendications précédentes, dans lequel les valeurs de gain produisent une quantité supérieure d'atténuation pour les voies à fréquence élevée que pour les voies à faible fréquence.

7. Système d'élimination de bruit selon la revendication 1, dans lequel ledit moyen (310, 810) permettant de produire une valeur de gain comporte en outre une pluralité de tables de gain, chaque table de gain présentant une valeur de gain de voie individuelle prédéterminée correspondant auxdites estimations d'énergie de voie individuelle, et un moyen de sélection de tables de gain permettant de choisir automatiquement une table de gain parmi la pluralité de tables de gain en fonction de niveau de bruit de fond global moyen dudit signal d'entrée.

8. Système d'élimination de bruit selon l'une quelconque des revendications précédentes, comportant en outre un moyen permettant de combiner ladite pluralité de signaux post-traités pour produire ledit signal de sortie dont on a supprimé le bruit.

9. Système d'élimination de bruit selon l'une quelconque des revendications précédentes dans lequel ledit moyen (220) permettant de générer des estimations comporte :

un moyen permettant de générer et de stocker une estimation de la densité spectrale de puissance du bruit de fond desdits signaux prétraités, ledit moyen de génération d'estimation de bruit de fond comportant un moyen permettant de modifier ladite estimation de bruit de fond en réponse à un paramètre temporel indiquant l'intervalle de temps s'étant écoulé depuis la modification de l'estimation de bruit de fond précédente ; et
un moyen permettant de générer une estimation du rapport signal/bruit (SNR) dans chaque voie individuelle en fonction desdites estimations de bruit de fond modifiées ;
dans lequel ledit moyen (310, 810) permettant de produire une valeur de gain pour chaque voie individuelle produit chaque valeur de gain en réponse auxdites estimations de SNR de voie.

10. Système d'élimination de bruit selon la revendication 9, dans lequel ledit moyen de modification de l'estimation de bruit de fond comporte un moyen permettant de produire ledit paramètre temporel, et un moyen permettant de comparer ledit paramètre temporel à un seuil temporel prédéterminé de sorte qu'une modification des estimations de bruit de fond est effectuée lorsque ledit paramètre temporel dépasse ledit seuil temporel.

11. Système d'élimination de bruit selon la revendication 10, dans lequel ledit moyen de modification de l'estimation de bruit de fond comporte en outre un moyen permettant de générer une estimation de l'énergie dans chaque voie individuelle, et un moyen permettant de produire un paramètre d'énergie multivoie en réponse à la valeur totale de toutes les estimations d'énergie de voie individuelle.

12. Système d'élimination de bruit selon la revendication 11, dans lequel ledit moyen de production de paramètre d'énergie multivoie permet des variations mineures des estimations d'énergie de voie individuelle de sorte que lesdites variations mineures n'affectent pas de façon considérable ledit paramètre d'énergie multivoie.

13. Système d'élimination de bruit selon la revendication 11 ou 12, dans lequel ledit moyen de modification des estimations de bruit de fond comporte en outre un moyen permettant de comparer ledit paramètre d'énergie multivoie à un seuil d'énergie prédéterminé de sorte qu'une modification des estimations de bruit de fond est effectuée lorsque ledit paramètre d'énergie multivoie est inférieur audit seuil d'énergie.

14. Système d'élimination de bruit selon la revendication 13, dans lequel ledit moyen de modification des estimations de bruit de fond modifie lesdites estimations de bruit de fond en réponse audit paramètre temporel quel que soit ledit paramètre d'énergie multivoie.

15. Système d'élimination de bruit selon l'une quelconque des revendications 11 à 14, dans lequel ledit paramètre d'énergie multivoie est généré en traduisant lesdites estimations de SNR de voie individuelle en les métriques de parole de voie individuelle, la somme des métriques de parole étant une mesure de toutes les caractéristiques ressemblant à de la parole de l'énergie dans toutes les voies.

16. Système d'élimination de bruit selon l'une quelconque des revendications précédentes, comprenant en outre :

un moyen (815, 820, 830) permettant de surveiller lesdites estimations d'énergie de voie et de distinguer des salves de bruit à bande étroite de l'énergie de parole et de l'énergie de bruit de fond, ce qui permet de produire un signal de modification (835) ;
un moyen (590) permettant de modifier de façon sélective lesdites estimations d'énergie de voie en réponse audit signal de modification (835) de sorte que les estimations d'énergie de voie représentant des salves de bruit à bande étroite soient modifiées ;
un moyen permettant de produire une valeur de gain pour chaque voie individuelle en réponse à chaque estimation d'énergie de voie modifiée ; et
un moyen permettant de modifier le gain de chaque signal prétraité de ladite pluralité de signaux prétraités en réponse auxdites valeurs de gain pour fournir une pluralité de signaux post-traités.

17. Système d'élimination de bruit selon la revendication 16, dans lequel ledit signal de modification indique le nombre total de voies individuelles présentant des estimations d'énergie dépassant un seuil d'énergie prédéterminé.

18. Système d'élimination de bruit selon la revendication 16 ou 17, dans lequel ledit moyen de modification des estimations d'énergie de voie comporte un moyen (830) permettant de comparer ledit signal de modification à un seuil

de compte prédéterminé de sorte qu'une modification des estimations d'énergie de voie est effectuée lorsque ledit nombre total de voies individuelles est inférieur audit seuil de compte.

19. Système d'élimination de bruit selon la revendication 16, 17 ou 18, dans lequel ledit moyen de modification de gain fournit une quantité maximum d'atténuation du signal prétraité dans une voie particulière ayant une estimation d'énergie de voie modifiée.

20. Procédé d'atténuation du bruit de fond provenant d'un signal d'entrée bruyant (205) pour produire un signal de sortie dont on a supprimé le bruit (265) dans un système d'élimination de bruit (800) comprenant les étapes consistant à :

séparer (850) le signal d'entrée en une pluralité de signaux prétraités représentant un nombre N de voies de fréquence choisies ;
générer une estimation de l'énergie dans chaque voie individuelle (853) ;
générer et stocker une estimation de la densité spectrale de la puissance du bruit de fond desdits signaux prétraités ; et
générer une estimation du rapport signal/bruit (SNR) dans chaque voie individuelle en fonction desdites estimations de bruit de fond et desdites estimations d'énergie de voie ;

le procédé étant caractérisé par les étapes consistant à :

produire (861, 862) une valeur de gain pour chaque voie individuelle en réponse auxdites estimations de SNR de voie, dans lequel lesdites valeurs de gain présentent un éventail de valeurs minimales ;

et dans lequel ladite étape de production des valeurs de gain comporte les étapes consistant à :

fournir un seuil de SNR prédéfini et comparer (902) lesdites estimations de SNR de voie audit seuil de SNR prédéfini de sorte que les voies ayant des estimations de SNR inférieures audit seuil de SNR produisent des valeurs de gain contenues dans ledit éventail minimal ; et
modifier (910) le gain de chaque signal prétraité de ladite pluralité de signaux prétraités en réponse auxdites valeurs de gain afin de fournir une pluralité de signaux post-traités.

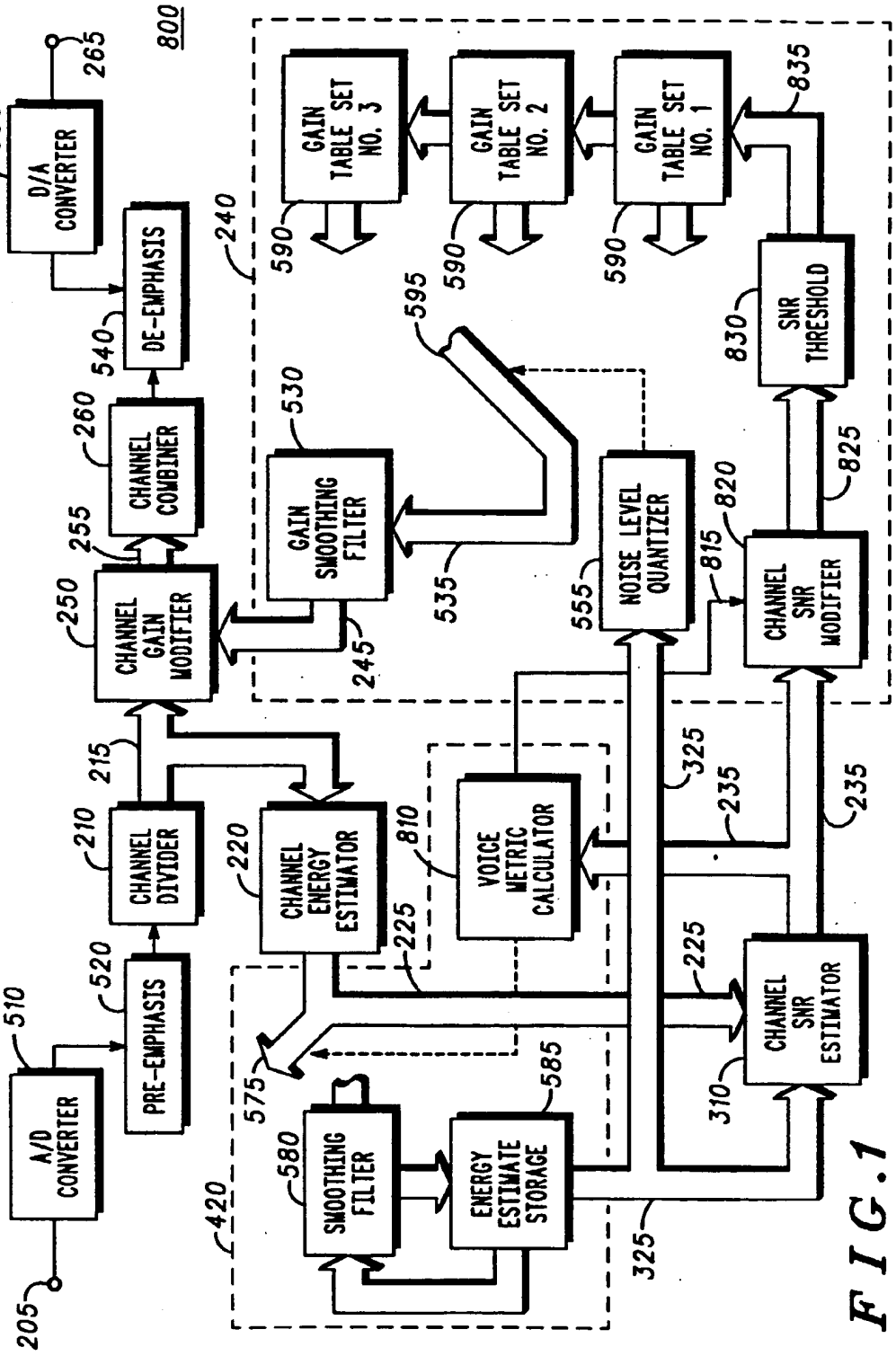

800
205
A/D CONVERTER
510
PRE-EMPHASIS
520
CHANNEL DIVIDER
210
215
CHANNEL GAIN MODIFIER
250
255
CHANNEL COMBINER
260
DE-EMPHASIS
540
D/A CONVERTER
550
265
CHANNEL ENERGY ESTIMATOR
220
225
VOICE METRIC CALCULATOR
810
GAIN SMOOTHING FILTER
530
535
595
245
240
GAIN TABLE SET NO. 3
GAIN TABLE SET NO. 2
GAIN TABLE SET NO. 1
590
835
NOISE LEVEL QUANTIZER
555
CHANNEL SNR MODIFIER
815
820
SNR THRESHOLD
830
825
CHANNEL SNR ESTIMATOR
310
325
235
SMOOTHING FILTER
580
ENERGY ESTIMATE STORAGE
585
575
420

FIG.1

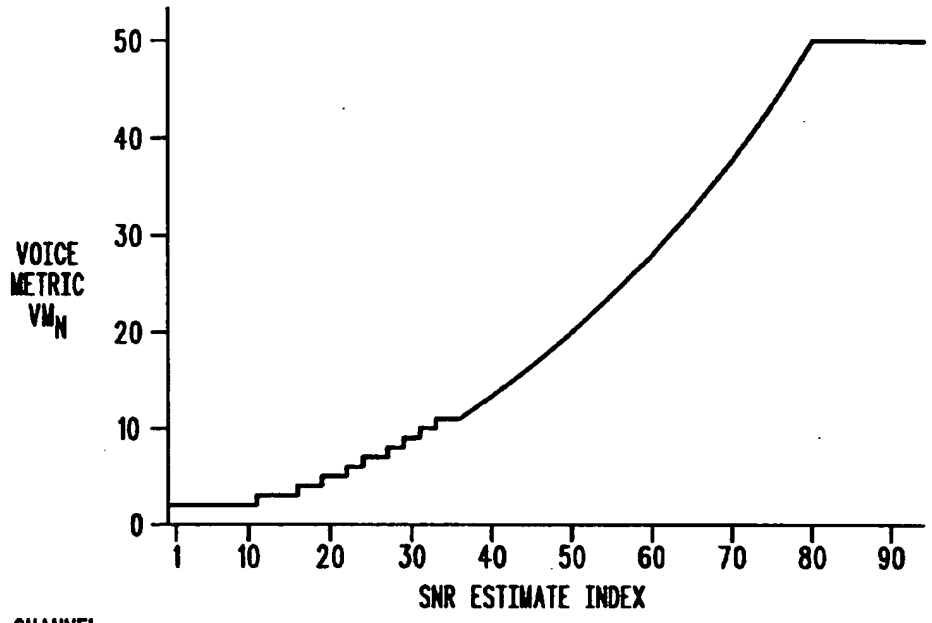

VOICE METRIC VM_N
50
40
30
20
10
0
1
10
20
30
40
50
60
70
80
90
SNR ESTIMATE INDEX

FIG.2

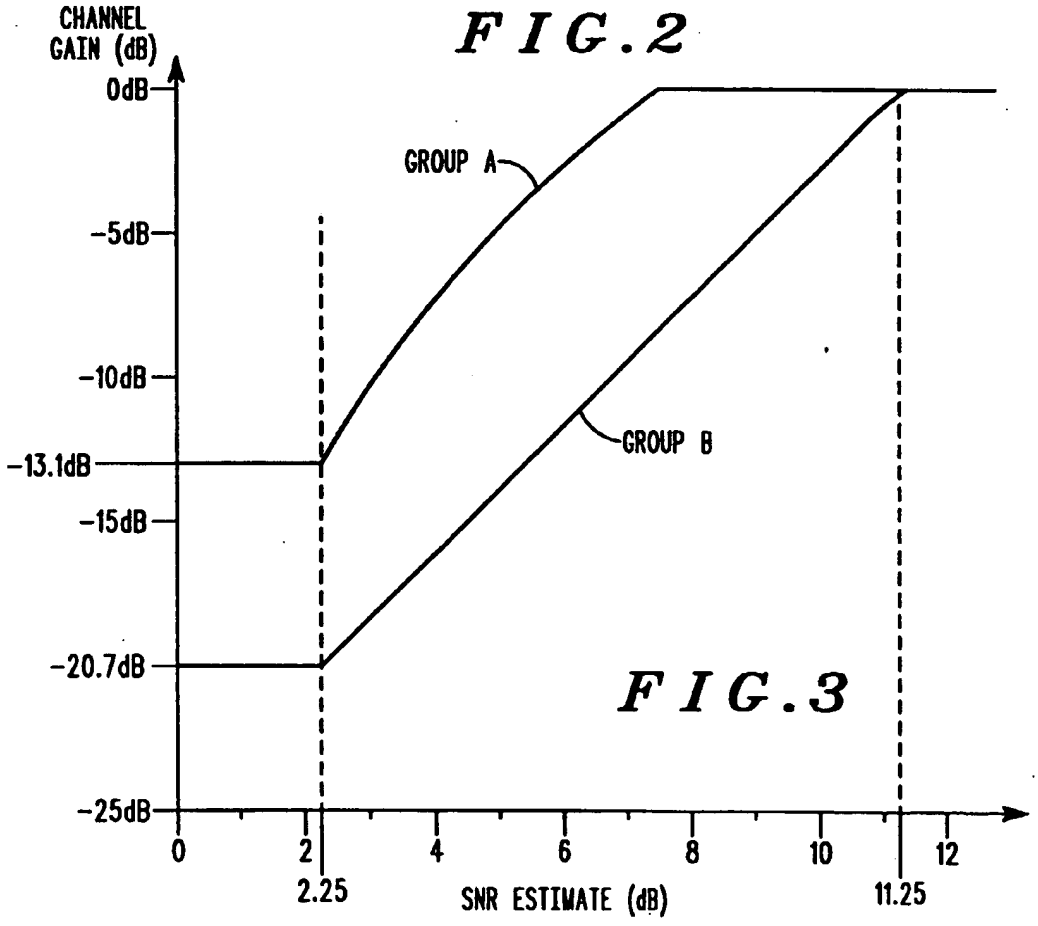
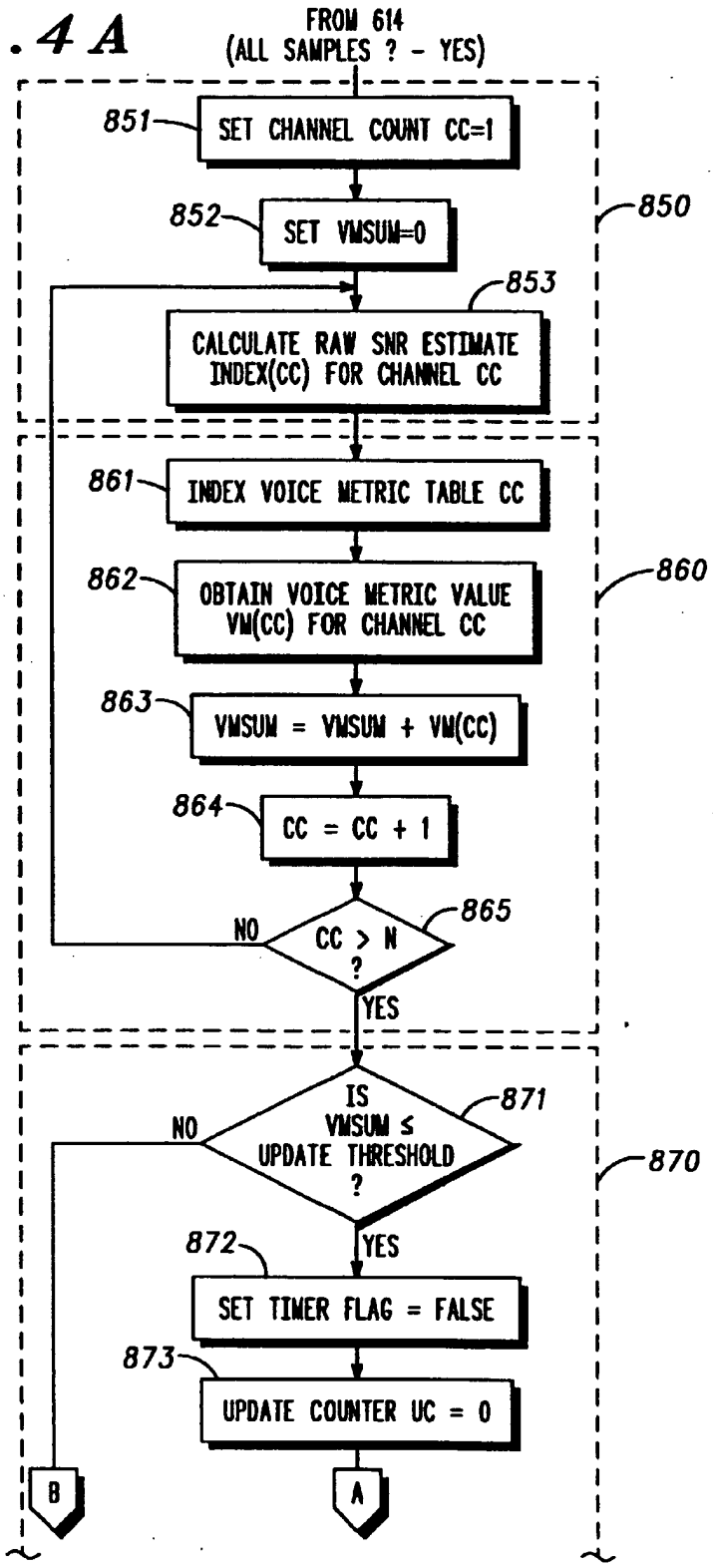

CHANNEL GAIN (dB)
0dB
-5dB
-10dB
-13.1dB
-15dB
-20.7dB
-25dB
GROUP A
GROUP B
0
2
2.25
4
6
8
10
11.25
12
SNR ESTIMATE (dB)

FIG.3

FIG. 4A


FROM 614
(ALL SAMPLES ? - YES)
851
SET CHANNEL COUNT CC=1
852
SET VMSUM=0
850
853
CALCULATE RAW SNR ESTIMATE INDEX(CC) FOR CHANNEL CC
861
INDEX VOICE METRIC TABLE CC
862
OBTAIN VOICE METRIC VALUE VM(CC) FOR CHANNEL CC
860
863
VMSUM = VMSUM + VM(CC)
864
CC = CC + 1
865
NO
CC > N ?
YES
871
IS VMSUM ≤ UPDATE THRESHOLD ?
NO
870
YES
872
SET TIMER FLAG = FALSE
873
UPDATE COUNTER UC = 0
B
A

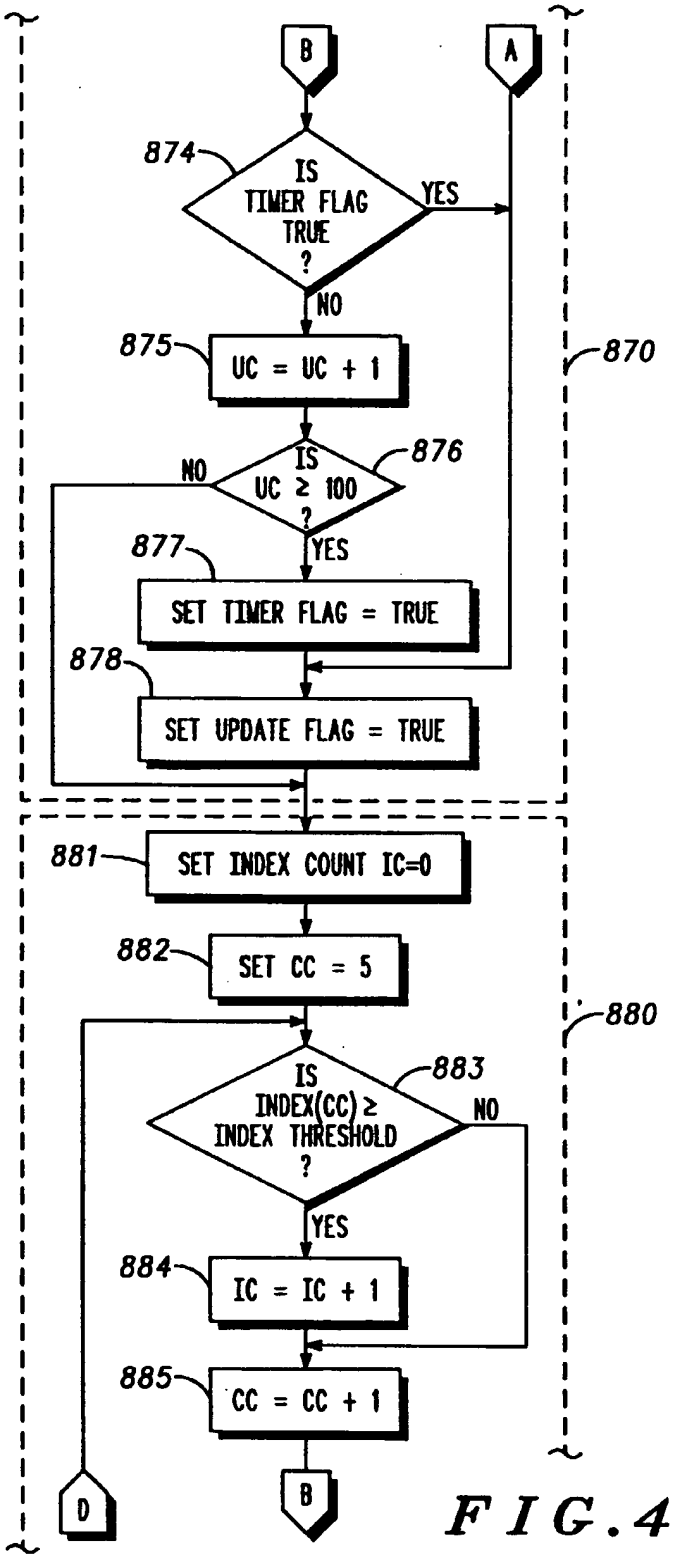
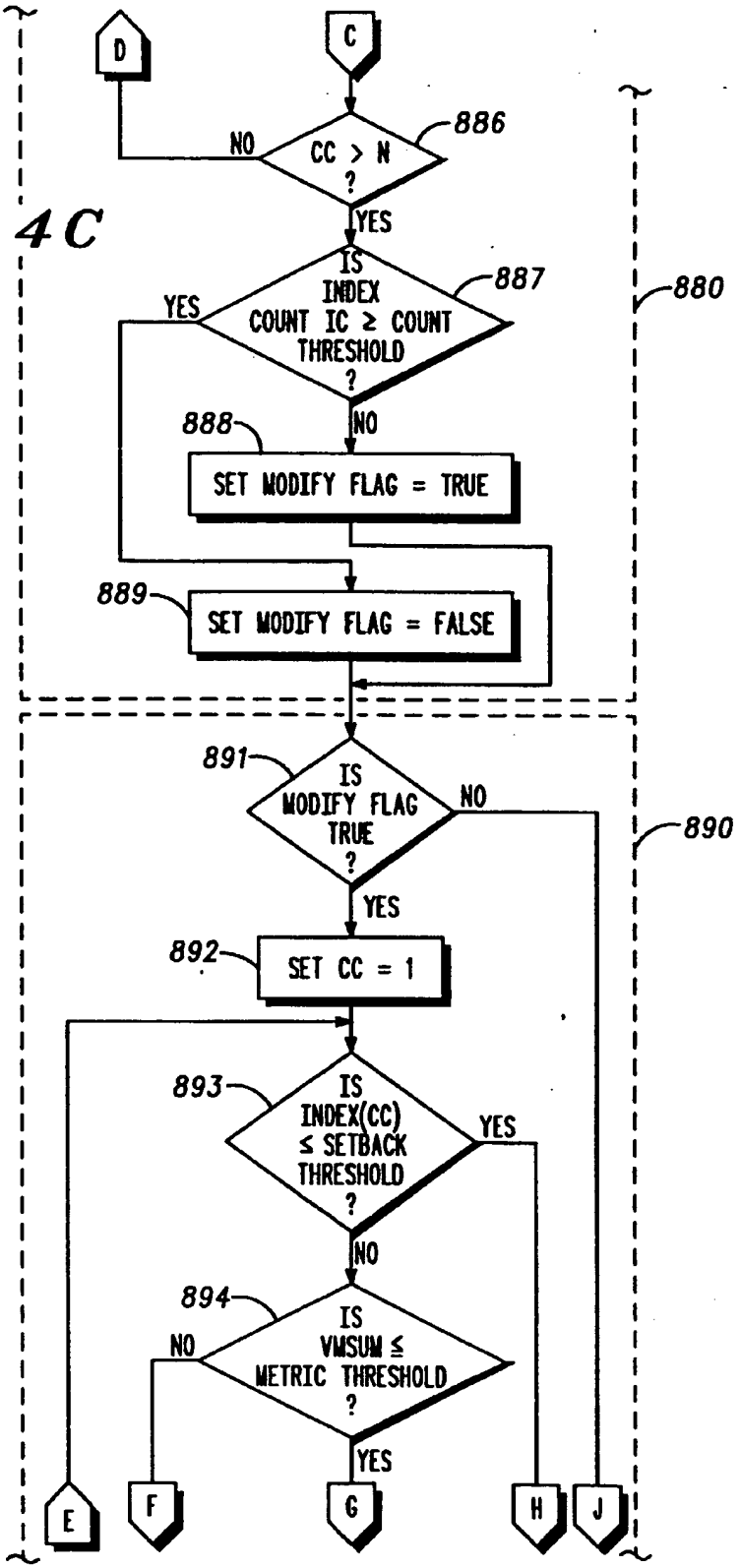

B
A
874
IS TIMER FLAG TRUE ?
YES
NO
875
UC = UC + 1
870
876
IS UC ≥ 100 ?
NO
YES
877
SET TIMER FLAG = TRUE
878
SET UPDATE FLAG = TRUE
881
SET INDEX COUNT IC=0
882
SET CC = 5
880
883
IS INDEX(CC) ≥ INDEX THRESHOLD ?
NO
YES
884
IC = IC + 1
885
CC = CC + 1
D
B

FIG.4B

FIG. 4C


D
C
886
NO
CC > N ?
YES
IS INDEX COUNT IC ≥ COUNT THRESHOLD ?
887
880
YES
NO
888
SET MODIFY FLAG = TRUE
889
SET MODIFY FLAG = FALSE
891
IS MODIFY FLAG TRUE ?
NO
890
YES
892
SET CC = 1
893
IS INDEX(CC) ≤ SETBACK THRESHOLD ?
YES
NO
894
IS VMSUM ≤ METRIC THRESHOLD ?
NO
YES
E
F
G
H
J

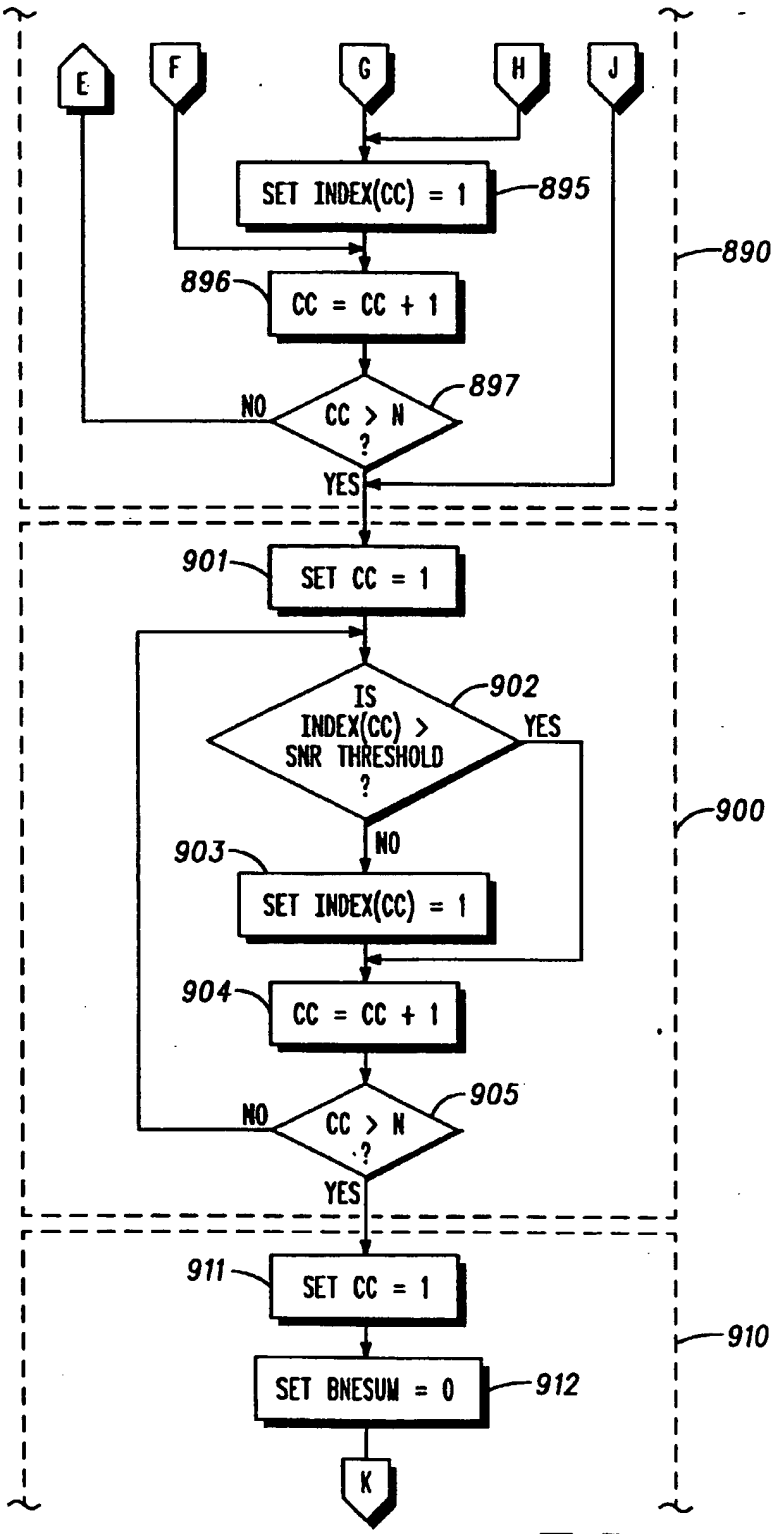

E
F
G
H
J
SET INDEX(CC) = 1
895
890
896
CC = CC + 1
897
NO
CC > N ?
YES
901
SET CC = 1
902
IS INDEX(CC) > SNR THRESHOLD ?
YES
NO
900
903
SET INDEX(CC) = 1
904
CC = CC + 1
905
NO
CC > N ?
YES
911
SET CC = 1
910
SET BNESUM = 0
912
K

FIG.4D

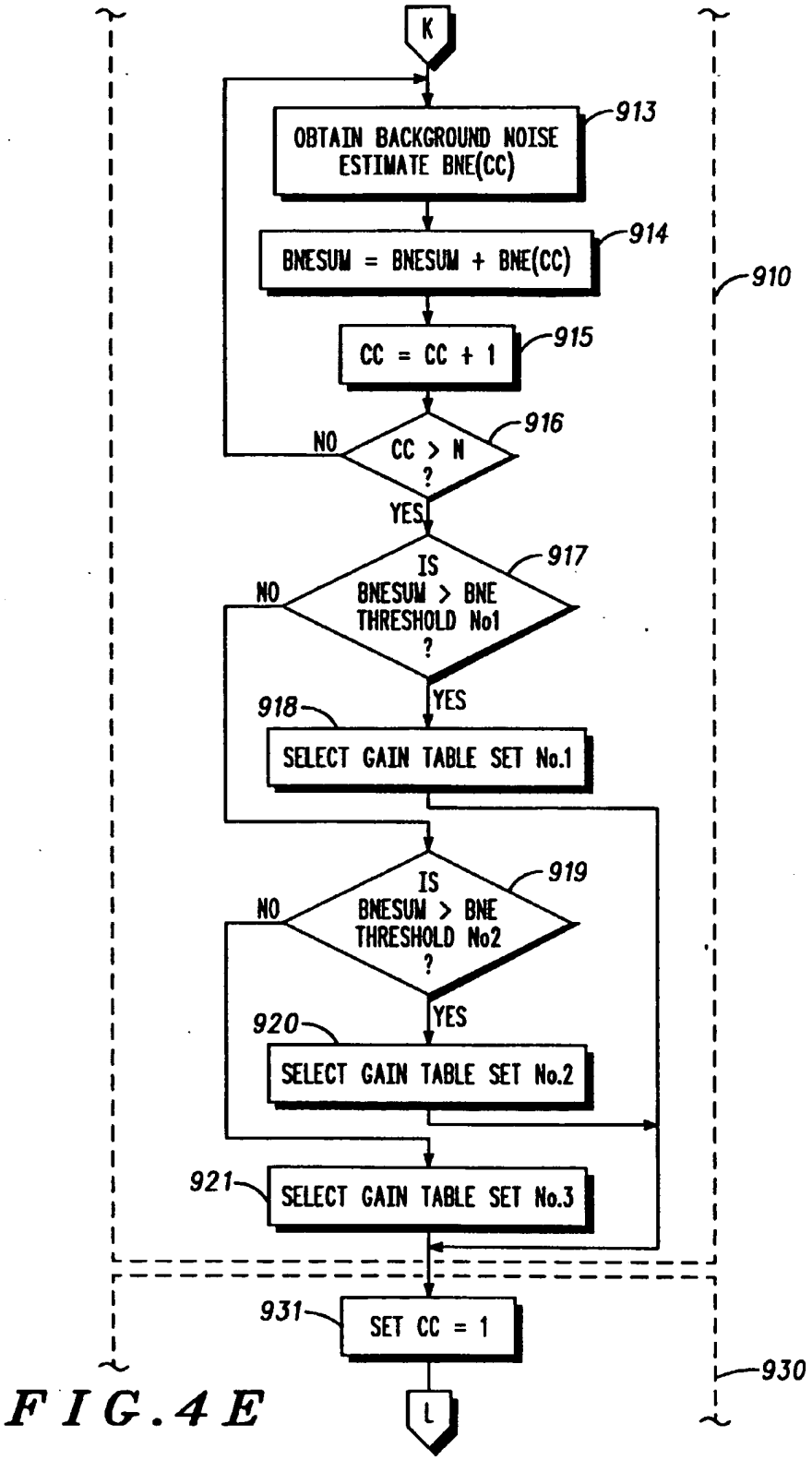

K
OBTAIN BACKGROUND NOISE ESTIMATE BNE(CC)
913
BNESUM = BNESUM + BNE(CC)
914
910
CC = CC + 1
915
NO
CC > N ?
916
YES
IS BNESUM > BNE THRESHOLD No1 ?
917
NO
YES
918
SELECT GAIN TABLE SET No.1
IS BNESUM > BNE THRESHOLD No2 ?
919
NO
YES
920
SELECT GAIN TABLE SET No.2
921
SELECT GAIN TABLE SET No.3
931
SET CC = 1
930
L

FIG.4E

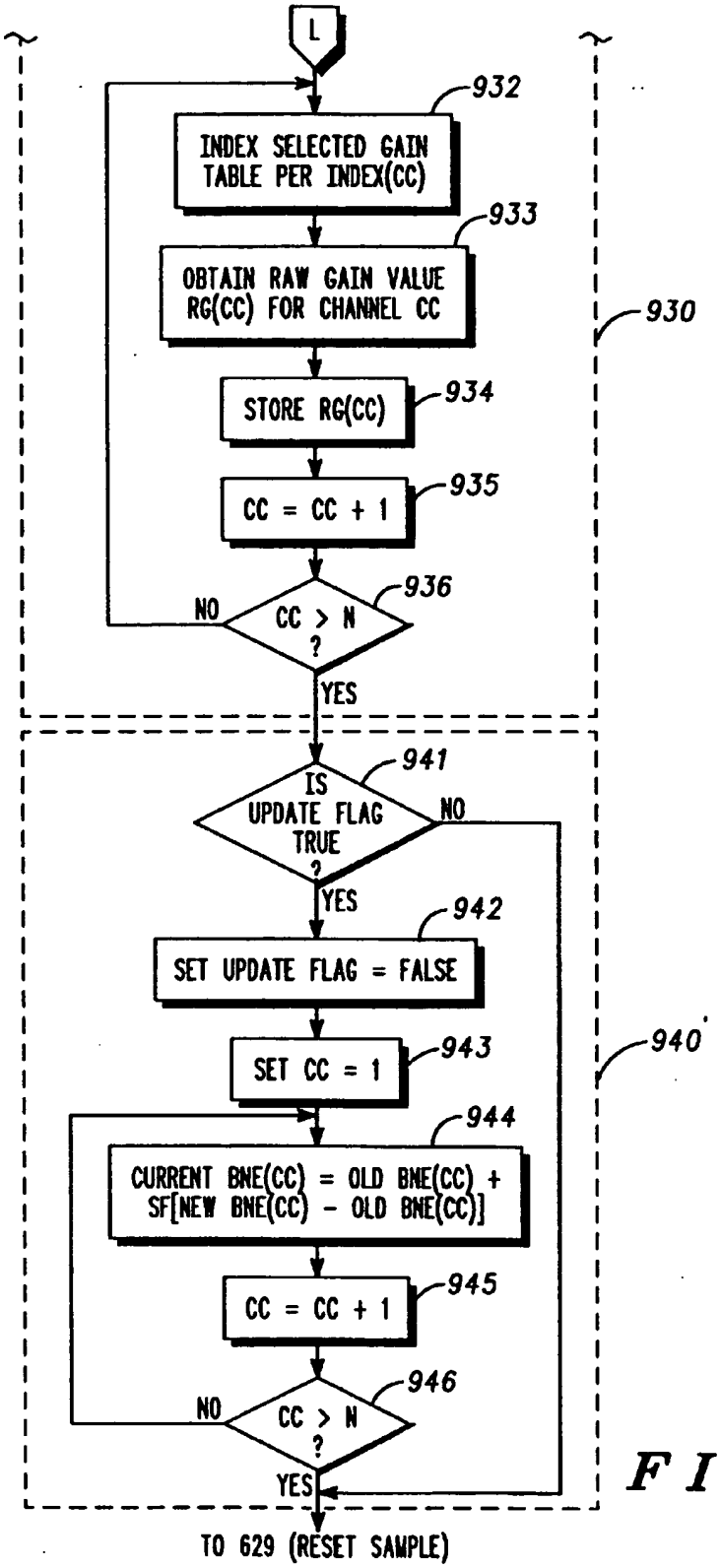

L
932
INDEX SELECTED GAIN TABLE PER INDEX(CC)
933
OBTAIN RAW GAIN VALUE RG(CC) FOR CHANNEL CC
930
STORE RG(CC)
934
CC = CC + 1
935
936
NO
CC > N ?
YES
941
IS UPDATE FLAG TRUE ?
NO
YES
942
SET UPDATE FLAG = FALSE
940
SET CC = 1
943
944
CURRENT BNE(CC) = OLD BNE(CC) + SF[NEW BNE(CC) - OLD BNE(CC)]
CC = CC + 1
945
946
NO
CC > N ?
YES
TO 629 (RESET SAMPLE)

FIG.4F

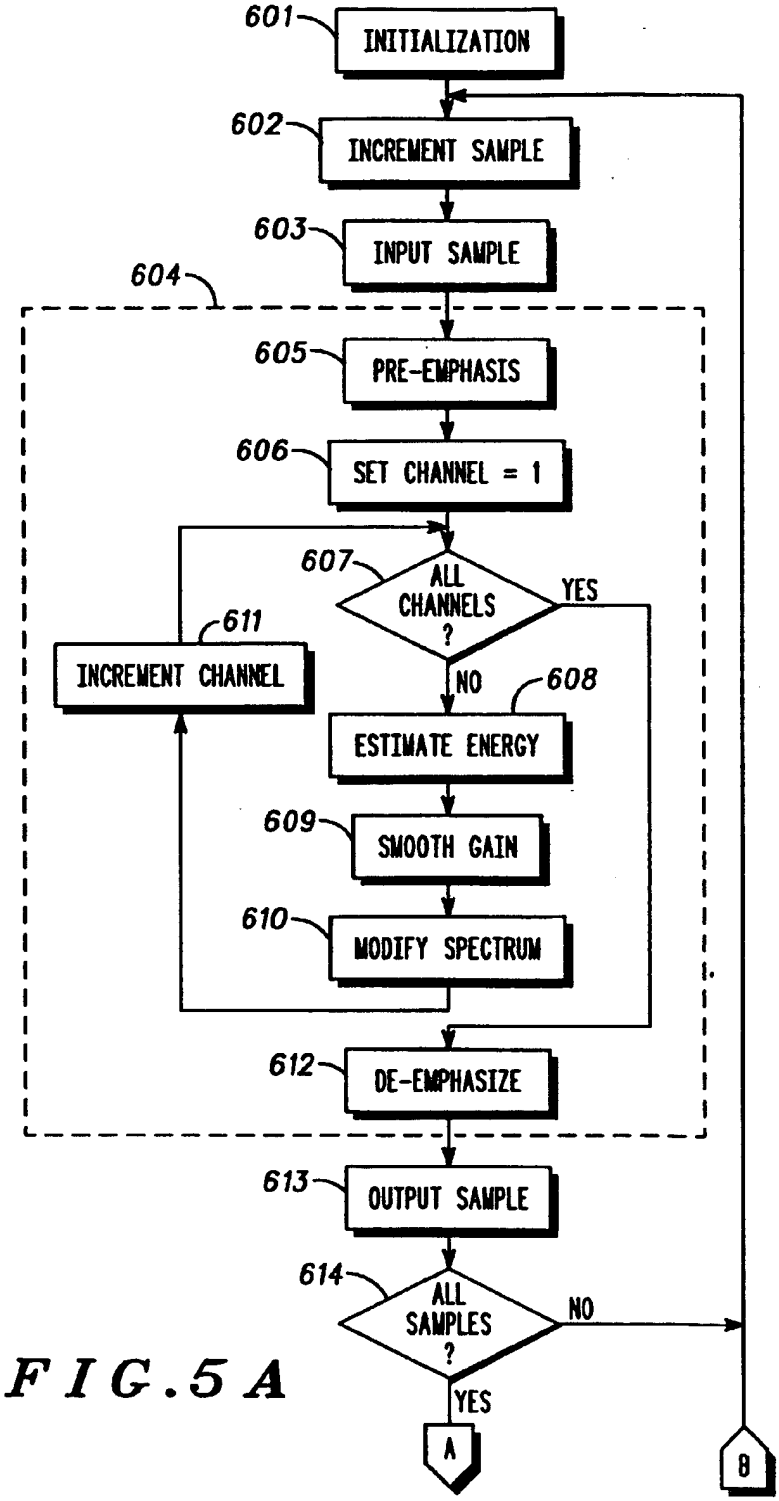
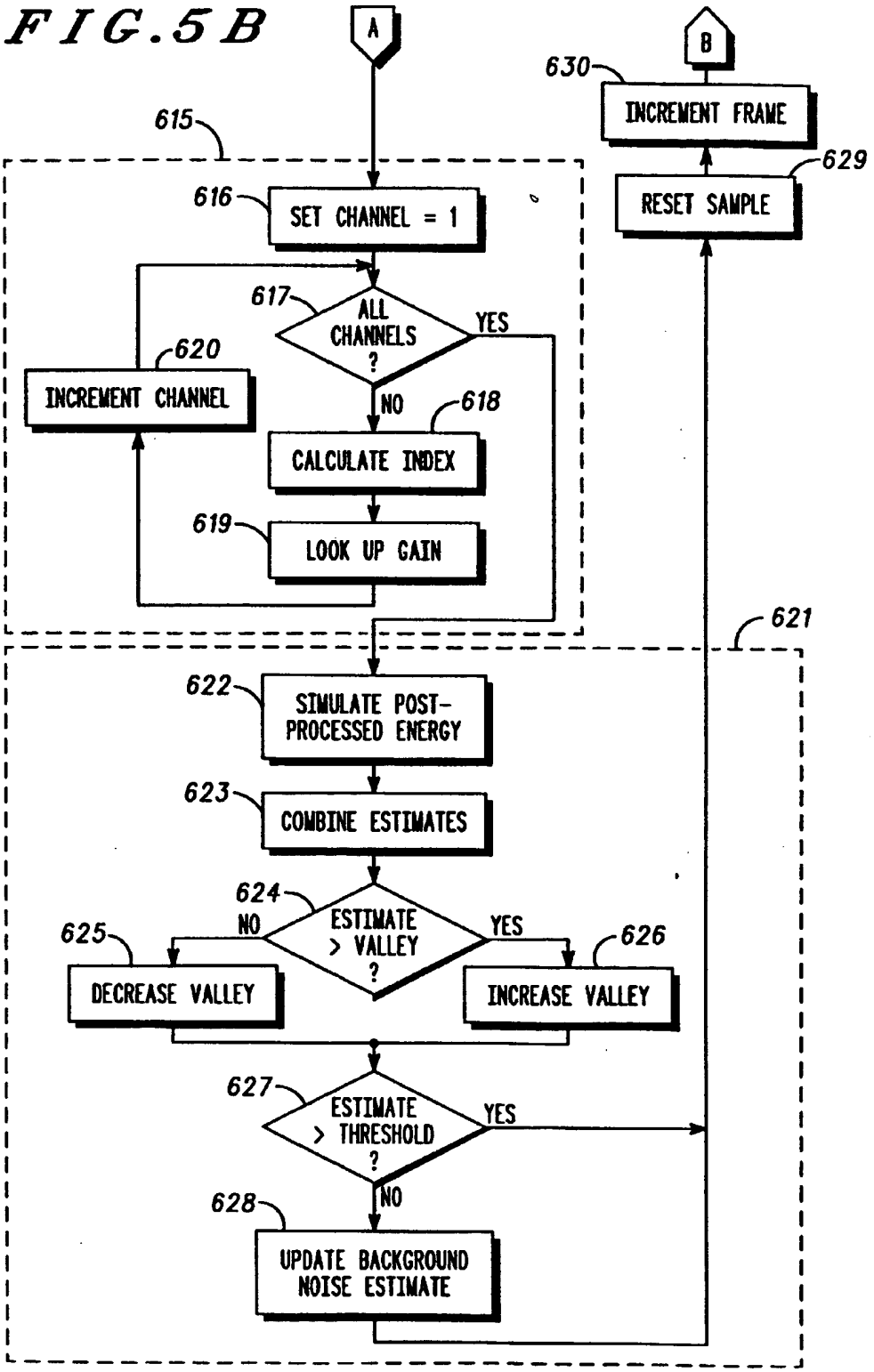

601
INITIALIZATION
602
INCREMENT SAMPLE
603
INPUT SAMPLE
604
605
PRE-EMPHASIS
606
SET CHANNEL = 1
607
ALL CHANNELS ?
YES
NO
611
INCREMENT CHANNEL
608
ESTIMATE ENERGY
609
SMOOTH GAIN
610
MODIFY SPECTRUM
612
DE-EMPHASIZE
613
OUTPUT SAMPLE
614
ALL SAMPLES ?
NO
YES
A
B

FIG.5A

FIG.5B


A
B
615
616
SET CHANNEL = 1
617
ALL CHANNELS ?
YES
NO
620
INCREMENT CHANNEL
618
CALCULATE INDEX
619
LOOK UP GAIN
630
INCREMENT FRAME
629
RESET SAMPLE
621
622
SIMULATE POST-PROCESSED ENERGY
623
COMBINE ESTIMATES
624
ESTIMATE > VALLEY ?
NO
YES
625
DECREASE VALLEY
626
INCREASE VALLEY
627
ESTIMATE > THRESHOLD ?
YES
NO
628
UPDATE BACKGROUND NOISE ESTIMATE

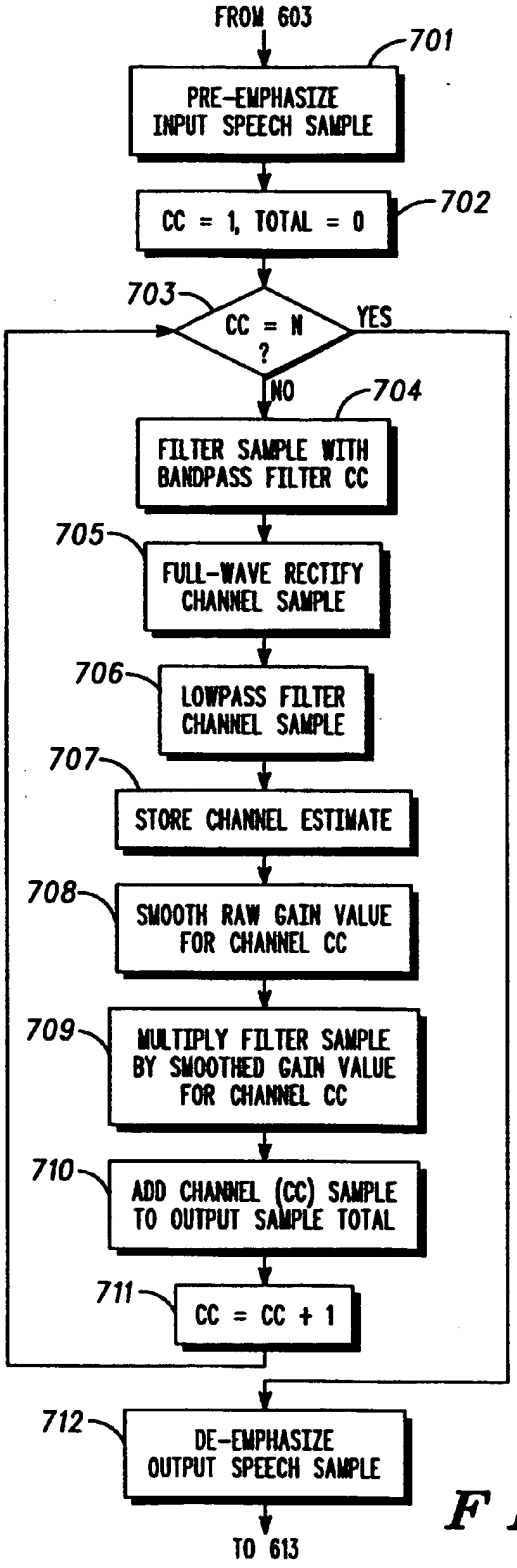

FROM 603
701
PRE-EMPHASIZE INPUT SPEECH SAMPLE
702
CC = 1, TOTAL = 0
703
CC = N ?
YES
NO
704
FILTER SAMPLE WITH BANDPASS FILTER CC
705
FULL-WAVE RECTIFY CHANNEL SAMPLE
706
LOWPASS FILTER CHANNEL SAMPLE
707
STORE CHANNEL ESTIMATE
708
SMOOTH RAW GAIN VALUE FOR CHANNEL CC
709
MULTIPLY FILTER SAMPLE BY SMOOTHED GAIN VALUE FOR CHANNEL CC
710
ADD CHANNEL (CC) SAMPLE TO OUTPUT SAMPLE TOTAL
711
CC = CC + 1
712
DE-EMPHASIZE OUTPUT SPEECH SAMPLE
TO 613

FIG.6A

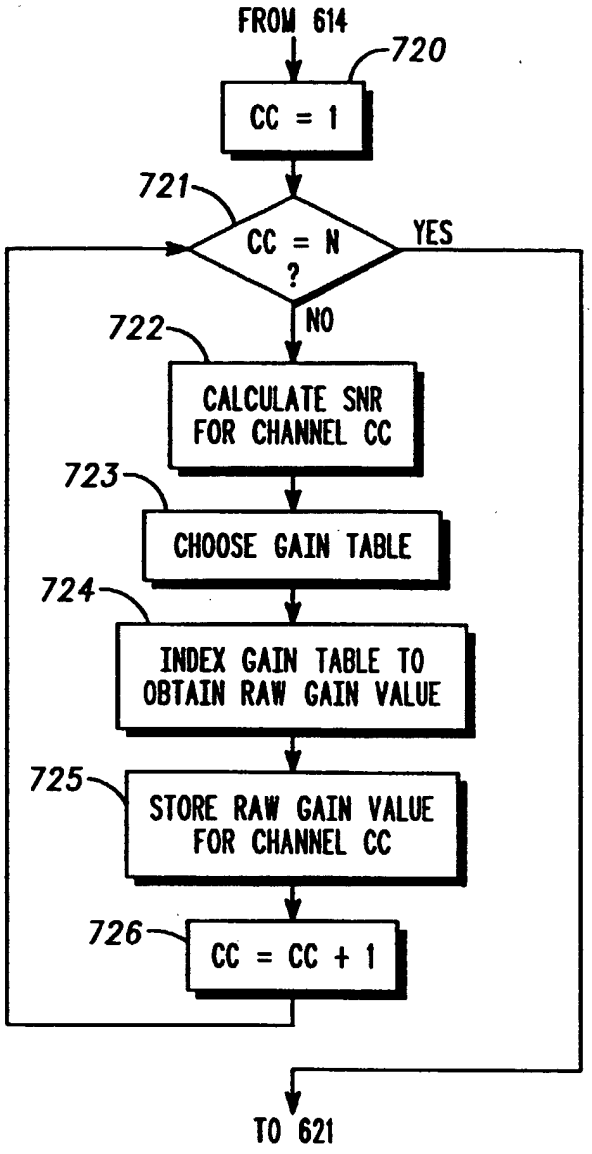

FROM 614
720
CC = 1
721
CC = N ?
YES
NO
722
CALCULATE SNR FOR CHANNEL CC
723
CHOOSE GAIN TABLE
724
INDEX GAIN TABLE TO OBTAIN RAW GAIN VALUE
725
STORE RAW GAIN VALUE FOR CHANNEL CC
726
CC = CC + 1
TO 621

FIG.6B

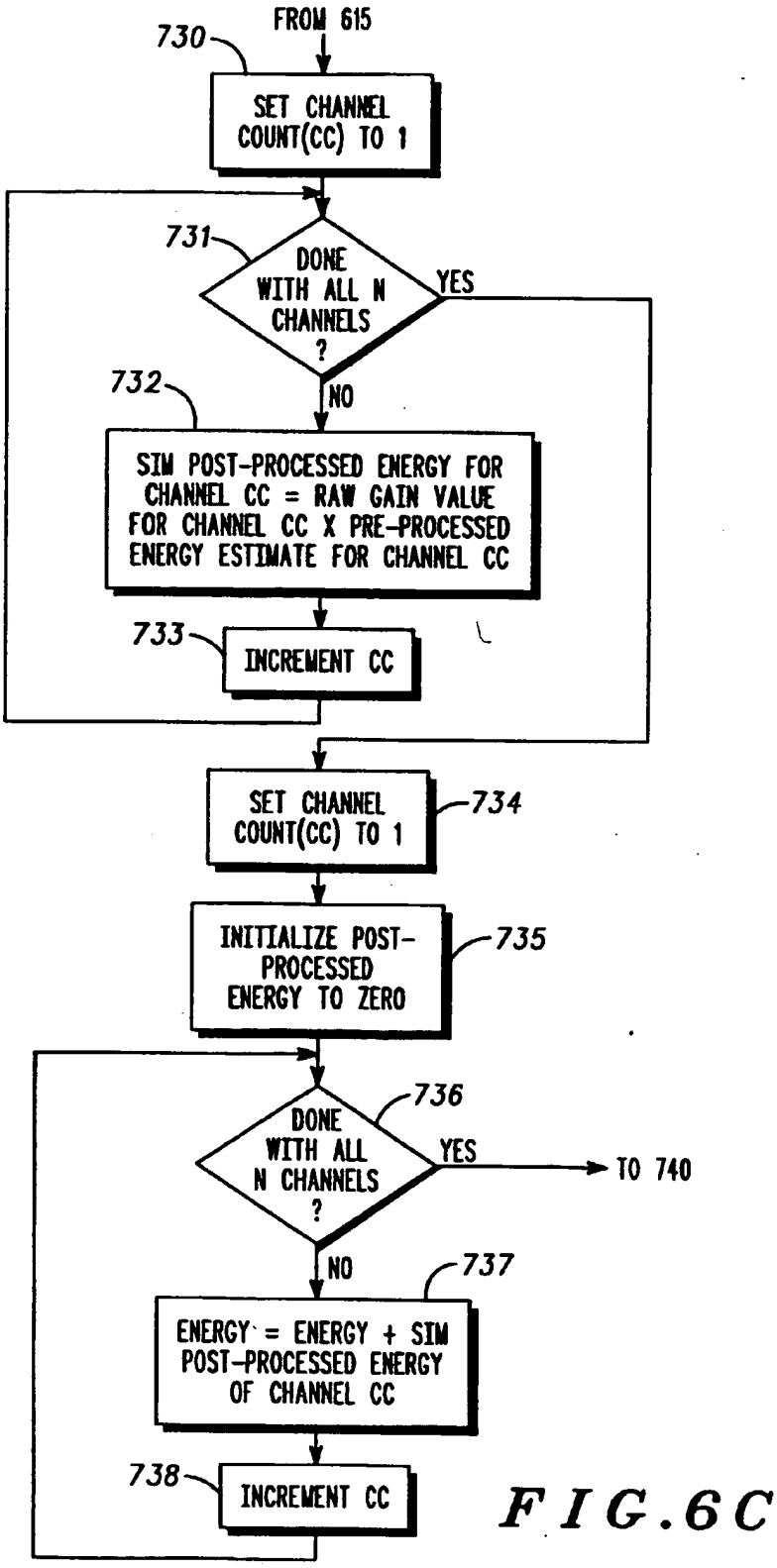

FROM 615
730
SET CHANNEL COUNT(CC) TO 1
731
DONE WITH ALL N CHANNELS ?
YES
NO
732
SIM POST-PROCESSED ENERGY FOR CHANNEL CC = RAW GAIN VALUE FOR CHANNEL CC X PRE-PROCESSED ENERGY ESTIMATE FOR CHANNEL CC
733
INCREMENT CC
SET CHANNEL COUNT(CC) TO 1
734
INITIALIZE POST-PROCESSED ENERGY TO ZERO
735
736
DONE WITH ALL N CHANNELS ?
YES
TO 740
NO
737
ENERGY = ENERGY + SIM POST-PROCESSED ENERGY OF CHANNEL CC
738
INCREMENT CC

FIG.6C

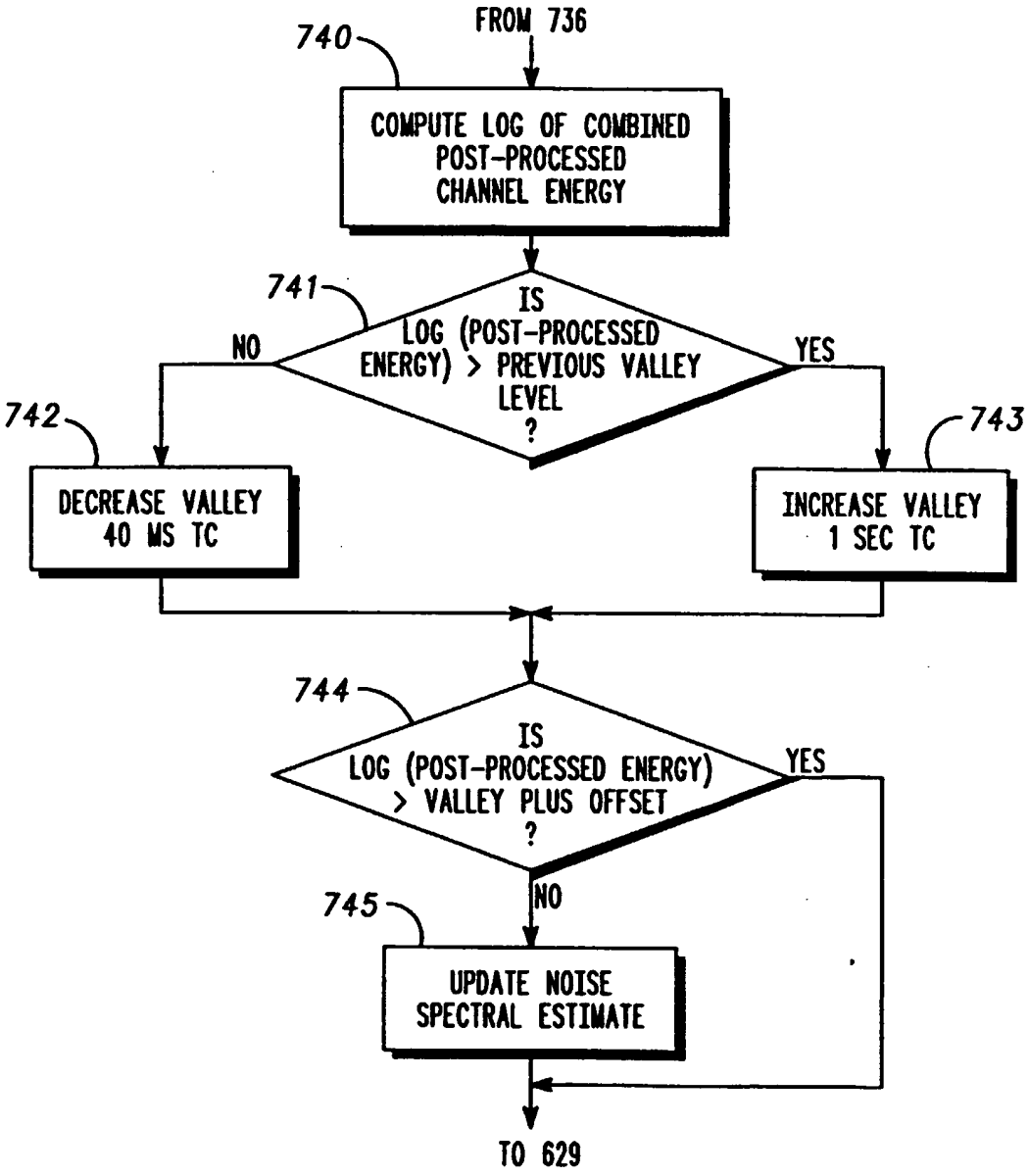

FROM 736
740
COMPUTE LOG OF COMBINED POST-PROCESSED CHANNEL ENERGY
741
IS LOG (POST-PROCESSED ENERGY) > PREVIOUS VALLEY LEVEL ?
NO
YES
742
DECREASE VALLEY 40 MS TC
743
INCREASE VALLEY 1 SEC TC
744
IS LOG (POST-PROCESSED ENERGY) > VALLEY PLUS OFFSET ?
YES
NO
745
UPDATE NOISE SPECTRAL ESTIMATE
TO 629

FIG.6D